2013

Molecular Functions of MLL PHD3 Binding to Its Ligands Cyp33 and H3K4Me3

Gayathree Raman
Loyola University Chicago

LOYOLA UNIVERSITY CHICAGO

MOLECULAR FUNCTIONS OF MLL PHD3 BINDING TO ITS LIGANDS CYP33 AND H3K4ME3

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

PROGRAM IN MOLECULAR BIOLOGY

BY

GAYATHREE RAMAN

CHICAGO, IL

MAY 2013

ACKNOWLEDGEMENTS

When I started my undergraduate studies, I was just another student with big dreams of higher studies. Through the years, several educators helped shape my pursuit of higher learning. I want to acknowledge the Loyola University Chicago Graduate School and the Molecular Biology Program for providing me the opportunity to undergo my graduate studies and helping my dream come true.

It's been great learning and conducting doctoral research under the mentorship of Manuel O. Diaz, M.D. Dr. Diaz encouraged me to be an independent researcher and at the same time provided timely guidance as needed. I am confident that the things I learned from Dr. Diaz will help me succeed in my scientific career.

Ute Osmers, Ph.D., is my go-to person whenever I had questions related to lab experiments or techniques or for trouble shooting. I thank her specifically for helping with mouse bone marrow isolation and methyl cellulose colony formation assay. Beyond all that, she is a good friend who is always there for me. Next, I want to thank and acknowledge Wei Wei for helping me understand critical steps in cloning and also for helping me with the cloning of MLL-N wt. I want to thank Peter Breslin, Jessica Solanki, Steven Poppen , Chris Bien and Carol Bier Lanning, M.D., for their valuable inputs and collaborations and for creating a congenial atmosphere in the laboratory.

I thank my doctoral committee members: Nancy Zeleznik-Le, Ph.D., Richard Schultz, Ph.D., Michael J. Thirman, M.D., and Chris Wiethoff, Ph.D., for their valuable guidance, motivation and support throughout my graduate research. I also thank all the faculty members of Molecular Biology and Bio-chemistry department who provided guidance whenever I reached out to them.

I thank Peggy Richied, Dona Buczek, Lorelei Hacholski and Janet Flores for patiently answering countless questions. I also thank Mitchell Denning, Ph.D., for his guidance throughout my doctoral program.

The valuable suggestions provided by the members of the Gene Regulation and Epigenetics Group greatly helped me in designing my experiments and also in trouble shooting. Specifically, I want to thank Nancy Zeleznik-Le, Ph.D. for being a great role model. I want to thank her lab members Shubin Zhang, M.D and Nick Aichelle for their help with retro-virus production, concentration and titration. I thank Jiwang Zhang, M.D and Dewen You, Ph.D. for their help with Flow Cytometry experiments.

Above all, I want to thank my family for being there whenever I needed them. I want to extend my sincere and heartfelt thanks to everyone who helped through my graduate studies.

To my dad

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Mixed Lineage Leukemia protein (MLL) is a key epigenetic regulator required for proper embryonic development, and fetal and adult hematopoiesis. It is a SET domain containing histone methyl transferase that trimethylates histone H3 on lysine 4 (H3K4Me3), a histone modification that correlates with active transcription. Like many chromatin-associated proteins, MLL has multiplePHD fingers. Previous studies from our laboratory have shown that the 3rd PHD finger of MLL interacts with the RNA Recognition Motif (RRM) of Cyp33. Cyp33 is a Cyclophilin with peptidyl prolyl isomerase (PPIase) activity. Over expression of Cyp33 results in transcriptional repression of MLL target genes. Thus Cyp33 switches MLL from a transcriptional activator to a repressor. In addition to binding to Cyp33, the MLL PHD3 also binds to H3K4Me3. Thus MLL is an H3K4Me3 "writer" with an embedded "reader" for the same mark. The aim of this study is to determine the biological function of MLL PHD3 binding to H3K4Me3 or Cyp33.

The *in vitro* binding studies performed by our collaborators, Sangho Park and John Bushweller, have shown that Cyp33 binding to MLL PHD3 reduces the binding affinity of the MLL PHD3 for H3K4Me3 by 5.7 fold. Therefore we sought to determine the effect of Cyp33 on H3K4 trimethylation and on transcription of MLL target genes. Cyp33 over expression resulted in a decrease in H3K4Me3 and H3Ac marks and increase in recruitment of JARID1A and JARID1B, the H3K4 specific demethylases, at the MLL

target gene promoters in a PPIase dependent manner. Also, shRNA- mediated down regulation of JARID1B resulted in transcriptional activation of a sub-set of MLL target genes. Thus we have identified JARID1B as transcriptional regulator of MLL target genes.

To determine the transcriptional outcome of the MLL PHD3 binding to H3K4Me3 or Cyp33, we abrogated either H3K4Me3 binding or Cyp33 binding to the MLL PHD3 by point mutations in the MLL-N coding sequence and assessed the effect of expressing the mutant proteins on MLL target gene expression. Over expressing wild type (wt) MLL-N resulted in increased expression of a sub-set of MLL target genes, increased recruitment of MLL-C and increased H3K4Me3 at MLL target gene promoters. Abrogating Cyp33 binding resulted in transcriptional activation of MLL target genes, however the effect was similar to that of wt MLL-N over expression. Unlike MLL-N wt , abrogation of H3K4Me3 binding to the MLL PHD3 did not result in transcriptional activation of MLL target genes, did not show an increase in MLL-C recruitment and did not increase H3K4Me3 levels. Also, abrogation of MLL PHD3 binding to H3K4Me3 resulted in decreased expansion of mouse hematopoietic progenitors in a methyl cellulose colony formation assay when compared to wt MLL-N. This indicates that the MLL PHD3 binding to H3K4Me3 contributes to the transcriptional activation of MLL target genes in hematopoietic progenitors. Thus, MLL PHD3 binding to H3K4Me3, contributes to the expansion of hematopoietic progenitors, demonstrating the biological significance. Taken together, the transcriptional outcome of MLL target genes is influenced by Cyp33

and H3K4Me3, the ligands of MLL PHD3. Therefore the MLL PHD3 serves as a ligand regulated molecular switch influencing the transcription of MLL target genes.

CHAPTER 1

INTRODUCTION

The MLL protein is a Trithorax group transcriptional regulator that is evolutionarily conserved throughout metazoans [1]. It belongs to the MLL family of SET domain containing histone methyl transferases which methylate histone H3 on lysine 4 [2–5]. The H3K4Me3 mark at gene promoters is associated with active transcription. In addition to its histone methyl transferase activity, MLL is a multi-domain containing protein. The domains of MLL allow it to associate with DNA, other chromatin regulators and histone modifiers, making it a versatile and a key epigenetic regulator.

MLL (MLL wt) is required for fetal and adult hematopoiesis and for embryonic development [6,7]. It is required for the maintenance of H3K4Me3 marks at *Hox gene* promoters [3]. *Hox* genes are a cluster of conserved, homeo-box containing transcription factors, responsible for the proper segmental identity during embryonic development and for the maintenance of tissue-specific developmental programs in adults [8]. In addition to its known role in embryonic development and hematopoiesis, MLL is also a cell cycle regulator. Cell cycle promoters of E2F target genes and cell cycle inhibitors like *CDKN1A* and *CDKN1B* are known targets of MLL [9–11].

Chromosomal rearrangements of the *MLL* gene are a frequent phenomenon in mixed lineage leukemia, hence the name *MLL* for the gene and protein [12,13].The MLL fusion proteins constitutively activate the transcription of a sub-set of MLL target genes

resulting in increased self-renewal of hematopoietic progenitors, keeping them in an undifferentiated state and eventually leading to leukemogenesis [14–16]. *MLL* rearranged leukemia are aggressive, are often associated with poor prognosis and the available treatments do not increase the survival of the patient significantly [17–19]. In *MLL* rearranged leukemia, only one of the *MLL* alleles is rearranged, the other allele is intact, and recent studies have demonstrated that MLL fusion protein requires the co-expression of the wild type MLL (MLL wt) for MLL fusion mediated transformation and leukemogenesis [20]. To design successful treatments for MLL rearranged leukemia one must know the relevant functions and contribution of MLL wt in leukemogenesis. MLL fusion proteins lacks PHD fingers, a bromo-domain, an activation domain and the SET domain, but retain the N-terminal region of MLL through the repression domain [21,22] . Studies from our laboratory (Chen et al., 2008) and from another group (Muntean et al., 2008) have shown that inclusion of the MLL PHD3 in a MLL fusion protein blocks MLL fusion mediated transformation of hematopoietic progenitors. Our present study focuses on understanding the role of MLL PHD3, and its ligands, in transcriptional regulation of MLL target genes.

Previous studies from our lab have shown that the RRM (RNA Recognition Motif) domain of Cyp33 is a ligand for MLL PHD3 [23]. Over expression of Cyp33 causes transcriptional repression of MLL target genes. Cyp33 belongs to the cyclophilin family of proteins. The RRM of Cyp33 binds AU rich RNA, is known to be present in spliceosomes and is involved in pre- mRNA splicing [24]. The cyclophilin domain of Cyp33 possesses Peptidyl Prolyl Isomerase (PPIase) activity which facilitates the cis-trans conformation change of prolyl peptide bonds [25]. [23] Double point mutations in the

Cyclophilin domain of Cyp33, R191A and F196A resulted in at least 99% reduction in the cis-trans proline isomerization (Wei, Breslin, Schultz, and Diaz, unpublished). The PPIase activity of Cyp33 is critical for the transcriptional repression of MLL target genes [26].

Over expression of Cyp33 augmented the recruitment of the transcriptional co-repressors HDAC1 and Bmi-1 [27]. MLL fusions lack PHD3 and are resistant to Cyp33 mediated repression of MLL target genes. The inclusion of the MLL PHD3 in MLL-ENL fusion proteins enabled the fusion proteins to interact with Cyp33 and recruit HDAC1 to the repression domain of MLL resulting in decreased expression of MLL target genes. As a consequence, the MLL fusion mediated transformation of mouse hematopoietic progenitors was blocked by inclusion of MLL PHD3 in MLLfusion[28].

Another ligand that binds MLL PHD3 is H3K4Me3. Recently, we and others have reported that the MLL PHD3 present in MLL-N, binds to H3K4Me3 [26,29,30]. The MLL PHD3 is a small domain of 60 amino acids. The Cyp33 binding site and H3K4Me3 binding sites are very close to each other. Hence binding of one ligand may affect the binding of the other. Furthermore, results from i*n vitro* binding studies performed by our collaborators (Park and Bushweller) have revealed that the presence of Cyp33 increased the Kd (dissociation constant) for MLL PHD3 binding to H3K4Me3 by 5.7 fold. Thus Cyp33 reduces the affinity of MLL PHD3 for H3K4Me3 [26]. If this is true *in vivo*, then an excess of Cyp33 would be predicted to affect the binding of MLL PHD3 to H3K4Me3. As a consequence, the H3K4Me3 marks would be exposed to histone demethylases resulting in demethylation of H3K4Me3 at MLL target genes promoters. The presence of the H3K4Me3 mark at MLL target gene promoters is associated with open chromatin and

transcriptional activation. Therefore a decrease in H3K4Me3 would result in decreased target gene expression. This hypothesis was tested by studying the effect of Cyp33 over expression on H3K4 trimethylation at MLL target genes. Cyp33 over expression results in increased recruitment of JARID1 histone demethylases and a corresponding decrease in H3K4Me3 levels at MLL target gene promoters. Thus this study has identified another mechanism of Cyp33-mediated transcriptional repression of MLL target genes.

JARID1 family proteins are histone H3K4Me3 specific demethylases [31–34]. Among the JARID1 family proteins, JARID1A and JARID1B have three PHD fingers. The PHD3 of JARID1A and JARID1B are known to bind to H3K4Me3[35]. They are H3K4Me3 "erasers" with an embedded "reader" for H3K4Me3 and therefore, can bind to H3K4Me3 and demethylate H3K4Me3 in adjacent histones or nucleosomes.

MLL is a "writer" with an embedded "reader". Nevertheless, the writer and reader functions are imparted by two separate protein fragments of MLL. The MLL protein is proteolytically cleaved into an N-terminal (MLL-N) and a C-terminal (MLL-C) fragments by the protease TASPASE-1 [36]. The PHD3 that binds to H3K4Me3 ("reader") is present on MLL-N, while the SET domain ("writer") is present on MLL-C. MLL-N and MLL-C interact directly through their FYRN and FYRC domains, and this interaction protects MLL-N from degradation by the E3 ubiquitin ligase SCF [Skp 37]. The nuclear localization signal in MLL-N is required for nuclear translocation of MLL-C. Thus the two different fragments require each other for functional holo-complex formation at the MLL target gene promoters [38]. Furthermore, the MLL-N and MLL-C interaction brings the reader and writer of H3K4Me3 together in the same complex. Hence we proposed that the MLL PHD3 binding to H3K4Me3 may help MLL-C to

methylate histone H3K4 on adjacent nucleosomes and also protect the methylated lysine 4 from histone demethylases, resulting in transcriptional activation of MLL target genes. Additionally, the activation domain of the MLL-C associates with histone acetyl transferases (HATs) like CBP and p300 [39]. The acetylation of histones opens the chromatin and allows the DNA to be accessible to transcription factors and the basal transcription machinery. The acetylation of histones is associated with transcriptional activation. MLL PHD3 binding to Cyp33 enhances the recruitment of transcriptional repressors like HDAC1 (Histone Deacetylase) and Bmi-1 leading to transcriptional repression of MLL target genes. For all the above mentioned reasons, we hypothesized that the MLL PHD3 binding to H3K4Me3 versus Cyp33 may have opposite effects on the transcriptional outcome for MLL target genes. Therefore, we wanted to determine the effect of abrogation of MLL PHD3 binding to H3K4Me3 or to Cyp33 on the transcription of MLL target genes in mouse embryonic fibroblasts (MEFs) and in primary mouse hematopoietic progenitors.

The point mutations that disrupt MLL PHD3 binding to H3K4Me3 or Cyp33 were introduced in the PHD3 encoding region of MLL-N. MLL-N wt or its mutants were over expressed in MEFs or in mouse hematopoietic progenitors and the effects on transcription of MLL target genes and on colony formation by hematopoietic progenitors was studied. Interestingly, over expression of MLL-N wt without over expressing the MLL-C resulted in increased H3K4Me3 marks, increased recruitment of MLL-C and increased transcriptional activation of a sub-set of MLL target genes. On the other hand, abrogation of MLL PHD3 binding to H3K4Me3 resulted in decreased H3K4Me3 marks and MLL-C recruitment at MLL target gene promoters, as well as decreased expression of a sub-set of

MLL target genes. More importantly, abrogation of MLL PHD3 binding to H3K4Me3 resulted in a drastic decrease in colony formation of primary mouse hematopoietic progenitors indicating a defect in their proliferation. The abrogation of MLL PHD3 binding to Cyp33 resulted in transcriptional activation of a sub-set of MLL target genes; however, the effects were similar to that of MLL-N wt over expression.

As mentioned earlier, recent studies have implicated the *Mll* wt allele in leukemogenesis and transformation of *Mll* fusion mediated leukemia. Milne et al., have shown that the MLL fusion protein binds to the *Hoxa9* promoter only when the promoter is already bound by the MLL wt protein. They identified the repression domain and PHD fingers as the minimal regions in MLL wt that contributes to the targeting of the MLL fusion to the *Hoxa9* promoter. Mutation of the PAF binding site in the repression domain and the H3K4Me3 binding site in the MLL PHD3 affected the MLL fusion's ability to bind to its target promoters[40]. Results from this study show that abrogating H3K4Me3 binding or Cyp33 binding to MLL PHD3 did not have any effect on MLL-AF9 mediated transformation of hematopoietic progenitors in a serial plating methyl cellulose colony formation assay. However, the effect of abrogation of H3K4Me3 binding of the MLL PHD3 on cell transformation by other MLL fusions has yet to be determined.

In conclusion, MLL PHD3 binding to H3K4Me3 contributes to H3K4 trimethylation, followed by transcriptional activation of a sub-set of MLL target genes and colony formation by hematopoietic progenitors. Additionally, we have identified a link between the JARID1 histone demethylases and Cyp33 mediated transcriptional repression of MLL target genes. Thus, the ligands of the MLL PHD3 influence the

transcription of MLL target genes and act as a molecular switch for transcriptional regulation of MLL target genes.

CHAPTER 2

REVIEW OF LITERATURE

Mixed Lineage Leukemia

Mixed Lineage leukemia is a distinct molecular sub-type of acute leukemia that occurs as a result of the chromosomal rearrangement of the 11q23 region. In these *MLL* rearrangements, the 11q23 chromosomal breakpoint region falls within the *MLL* gene and since this cytogenetic abnormality is found in both acute myeloid or acute lymphoid leukemia, often with lineage infidelity (expression of some markers of the opposite lineage), the gene is named after the disease [12]. The translocation of *MLL* with another gene results in an in-frame fusion gene that encodes for the leukemogenic MLL fusion protein. So far, there are 100 different recurrent *MLL* translocations reported, with at least 60 different partner genes [41]. MLL-rearranged leukemia accounts for 70% of all ALL and 35-50% of all AML in infants. In older children and adults it accounts for about 10% of all leukemia. Patients with MLL- rearranged leukemia show poor prognosis. The more common fusion partners of MLL are AF4, AF9, ENL, AF10, AF6, ELL, AF17, and SEPT6 which altogether account for 90% of MLL-rearranged leukemia [42].

Besides chromosomal rearrangement, gene amplification and partial tandem duplication (PTD) of *MLL* are other mechanisms of deregulation of MLL that occur frequently in acute leukemia. About 3- 20 copies of the entire *MLL* gene are found in the leukemic cells of older patients with acute leukemia or myelodysplatic syndrome [43]. The

exons 11-5 or 12-5 of the *MLL* gene are duplicated in the PTD of *MLL*. *MLL* PTD retains the TASPASE1 cleavage site and C-terminal histone methyl transferase activity. About 4-7% of AML patients have *MLL* PTD [44].

Domains of MLL Protein

MLL is a very large, 431 kDa protein with many different identified domains that mediate protein-DNA, protein-protein, or protein-RNA interactions. The MLL protein is proteolytically cleaved into an N- (MLL-N) and a C-terminal (MLL-C) fragment by the protease TASPASE1 [36]. The MLL-N fragment has a Menin binding region, 3 AT hooks, a repression domain, 4 PHD fingers, an atypical bromo domain and a FYRN (Phenylalanine and Tyrosine Rich N-terminus) domain. The MLL-C fragment has the transcriptional activation domain, a FYRC (Phenylalanine and Tyrosine Rich C-terminus domain). The chromosomal break point region is just before PHD1[12]. Hence the fusion protein contains a portion of MLL N-terminus through the repression domain fused in-frame with a C- terminal partner protein fragment. The rest of the N-terminus (from the PHD finger region to the TASPASE1 cleavage site) and the whole of MLL-C fragment is deleted in the MLL fusion protein as shown in Figure 1. Nevertheless, the deleted regions are frequently, but not always, represented in the reciprocal translocation product.

The first 43 amino acids of the N-terminus of MLL contains the Menin binding site[45]. Menin enhances the localization of MLL to its target promoters, is required for MLL's histone methyl transferase activity and thus for the MLL target gene expression [46,47]. Furthermore, Menin brings in other oncogenic factors like LEDGF and c-Myb to MLL and is essential for MLL fusion mediated leukemogenesis [48,49].

The MLL AT hooks, similar to the ones present in the HMG group proteins, bind AT-rich DNA sequences. However, the AT hooks of MLL do not recognize DNA in a sequence specific manner. Instead, they recognize and bind to cruciform DNA structures [21]. This region of MLL also is also a mediator of an interaction with the SET onco-protein, and with GADD34 [50,51]. The first 1,100 amino acids of MLL-N are required for MLL localization in a punctate distribution in the cell nucleus. Specifically, the SNL1 and SNL2 regions present in the N-terminal region of MLL contribute to the sub-nuclear localization of MLL [52].

The repression domain of MLL recruits transcriptional co-repressors like HDAC1, hPC2, Bmi-1 and CtBP [27]. Part of the MLL repression domain encodes the CXXC domain which binds to non-methylated CpG on the DNA [53]. Through this property, MLL protects CpG dinucleotide sequences from DNA methylation [54]. Recent studies have shown that the repression domain of MLL recruits PAF1, a member of the PAF-C transcriptional elongation complex, to MLL target gene promoters and is required for the MLL fusion mediated leukemogenesis and transformation [55].

There are three tandem PHD fingers and an atypical extended PHD finger in MLL. This region of MLL is highly conserved from *Drosophila* to humans. The PHD fingers of MLL homodimerize; however, the significance of homo-dimerization is yet to be determined [23]. The first and fourth PHD fingers of MLL contribute to the interaction between the MLL-N and MLL-C fragments [38]. A recent study has shown that the second PHD finger of MLL has an E3 ligase activity and requires the E2 ubiquitin conjugating enzyme CDC34 for its activity. Histones H3 and H4 were identified substrates of

ubiquitination *in vitro*. The PHD2 E3 ligase activity negatively regulates MLL target gene expression [56].

The third PHD finger of MLL is the region of interest of this study. Yeast two hybrid studies performed in our laboratory identified Cyp33 as a ligand for the MLL PHD3. Cyp33 over expression augmented the recruitment of the transcriptional repressors like HDAC1 and Bmi-1, to the repression domain and resulted in transcriptional repression of MLL target genes [23]. The PHD3 of Trithorax (TRX), the *Drosophila* homologue of MLL, also binds to *Drosophila* Cyp33 [57]. All the PHD fingers are deleted in leukemogenic MLL fusions. The insertion of the MLL PHD3 in a MLL-ENL fusion protein, enabled the fusion protein to interact with Cyp33 and recruit HDAC to the repression domain of MLL, resulting in decreased expression of MLL target genes [28]. Thus , the wild type MLL responds to transcriptional activation or transcription repression cues and switches on or off target genes expression. But the MLL fusions are resistant to the mechanisms of transcription repression and constitutively activate the target genes. Similar to other PHD finger containing chromatin-associated proteins, PHD3 of MLL binds to trimethylated H3K4 [26,29,30]. The binding of MLL PHD3 may contribute to the transcriptional activation of MLL target genes. The MLL fusion protein binds to the target promoters that are already bound by wild type MLL. MLL PHD3 binding to H3K4Me3 and PAF1 binding to the repression domain cooperate to recruit MLL wt to the target promoter [40].

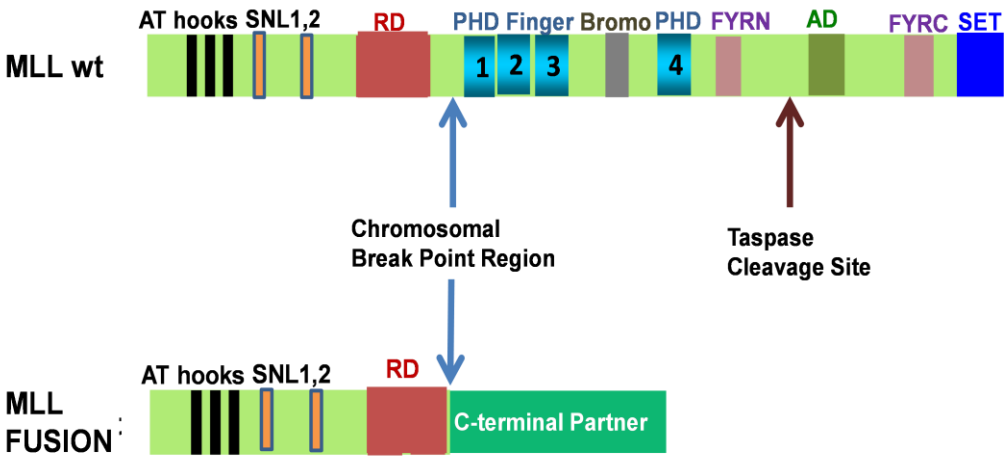
A short amino acid sequence between the PHD4 and the bromo-domain of MLL is required for interaction with HCF1, a cofactor of E2F1. This interaction recruits MLL to E2F responsive promoters [9]. The bromo-domain of MLL is an atypical one and does

not bind acetylated lysine, however its interaction with the MLL PHD3 increases the affinity of the latter for H3K4Me3[29]. The FYRN domain of MLL is a phenyl alanine and tyrosine rich region in the N-terminus which interacts with the FYRC region present in the C-terminus, thus tethering the MLL-N and MLL-C fragments together (Figure 2)[36].

The transcriptional activation domain of MLL-C binds to histone acetyl transferases like p300 and CBP [39]. The region adjacent to the activation domain interacts with the H4K16 acetyl transferase MOF and this interaction contributes to the transcriptional activation of *HOX* genes [58]. The SET domain of MLL is the catalytic domain that possesses the histone methyl transferase activity. The SET domain requires other core components of the COMPASS-like complex, RbBP5, ASH2L, WDR5 and DPY-30 to trimethylate histone H3 on lysine 4 [59].

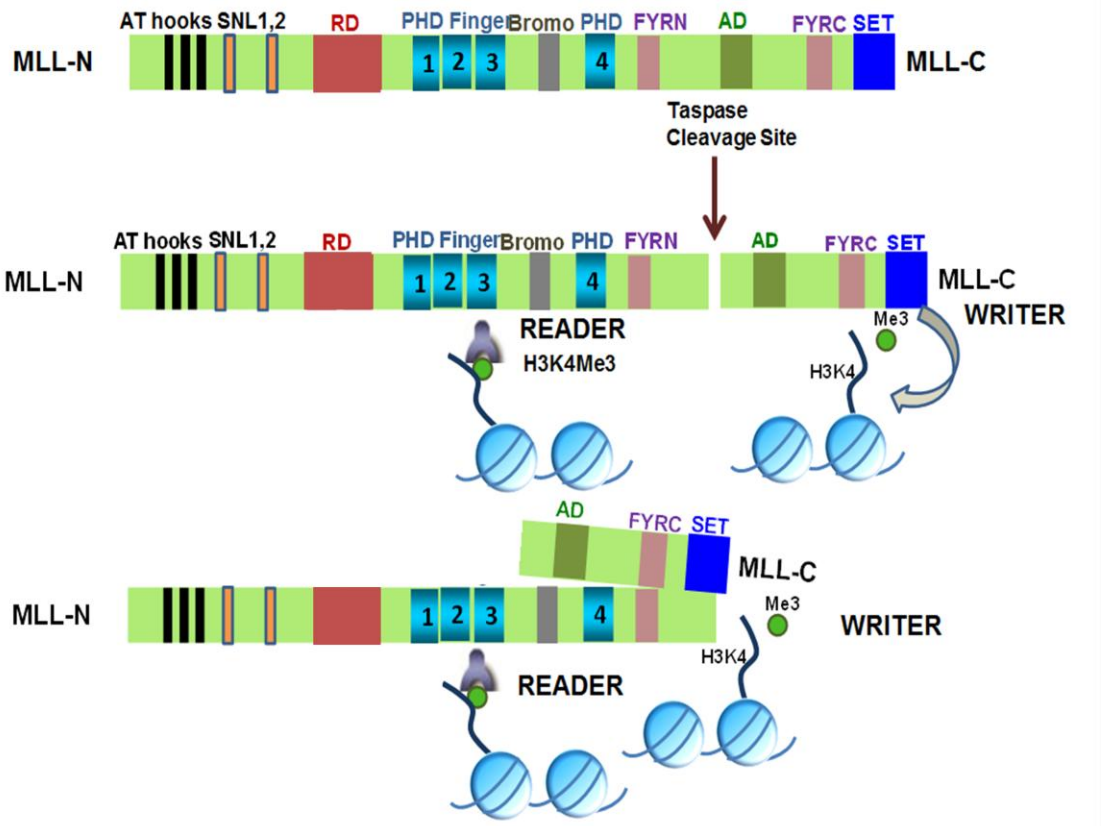
The SET domain has been reported to play a role in homo-dimerization and single stranded DNA and RNA binding [60,61]. The domains of MLL act as platforms for both homo and hetero-molecular interactions empowering MLL to be a key epigenetic regulator.

Figure 1: Illustration of domains of MLL wt and MLL fusion proteins.



MLL wt is proteolytically cleaved into MLL-N and MLL-C fragments by TASPASE1. Both MLL wt and the fusion have AT hooks, the Sub-Nuclear Localization (SNL) 1 and 2, repression domain (RD). The chromosomal break point region is just before or within the PHD1. Hence the region including the PHD fingers, the bromo-domain, the FYRN domain, the Activation domain (AD) and the SET domain is not present in the fusion proteins. MLL wt is a writer with an embedded reader. The SET domain is a histone methyl transferase that trimethylates histone H3 on lysine 4. The PHD3 is the reader that recognizes H3K4Me3 marks.

Figure 2: The hetero-dimerization of MLL-N and MLL-C allows the writer and reader to be present in the same complex.



MLL is a writer with an embedded reader for H3K4Me3. The SET domain of MLL is a histone methyl transferase that trimethylates histone H3 on lysine 4, hence it is a writer. The PHD3 of MLL binds to H3K4ME3; hence it is a reader of H3K4Me3. Though the writer and reader functions are present in the same protein, due to proteolytic cleavage by TASPASE1 they are separated in two different polypeptides. The proteolytic cleavage of MLL by TASPASE1, results in a 320 kDa MLL-N fragment and a 180 kDa MLL-C fragment. The MLL-N and MLL-C interact through their FYRN and FYRC domains. This interaction protects MLL-N from degradation and also is required for nuclear localization of MLL-C. Also, allows the writer and reader to be present in the same complex.

Post Translational Regulation of MLL

MLL is proteolytically cleaved by TASPASE1 and this cleavage is essential for the histone methyl transferase activity of MLL and transcription regulation of MLL target genes. The nuclear localization signal in MLL-N is required for nuclear translocation of MLL-C. MLL-N is targeted for degradation during the S and M phases of the cell cycle by SCF ^{Skp} and APC ^{Cdc20} E3 ligases [37]. The interaction of MLL-N and MLL-C stabilizes MLL-N and protects it from degradation. The MLL-C protein that is not bound to MLL-N cannot enter the nucleus and is targeted for proteasome-mediated degradation in the cytoplasm. Thus the two fragments require each other for formation of the functional enzymatically active holo-complex at target gene promoters [38]. Furthermore, MLL-N and MLL-C interaction brings the reader and writer of H3K4Me3 together in the same complex as shown in Figure 2. MLL fusions do not have the TASPASE1 cleavage site; hence they are resistant to cell cycle specific degradation. MLL is phosphorylated at Ser 516 by ATR. This phosphorylation event is a response to DNA damage and is required for correct execution of the S-phase check point. Phosphorylation at Ser 516 disrupts the interaction with SCF ^{Skp} and protects MLL from degradation. The accumulated MLL then methylates the histone H3K4 at the late replicating origins, and inhibits CDC45 loading, thus delaying DNA replication. The MLL fusion gets phosphorylated at Ser 516 but lacks the SET domain, hence resulting in a dysregulated S phase check point [62].

Molecular Complexes Containing MLL

The very first immunoaffinity purification of MLL complex was performed by Nakamura et al. That study identified MLL is present in complex with 28 other proteins. Other notable proteins of this complex are global transcription regulators that belong to

the core transcriptional machinery like TFIID and TBP; ATP-dependent nucleosome remodeling proteins like SWI/SNF, NuRD (HDAC), hSNF2H and Sin3A (HDAC); and other core components of the yeast Set1 histone methyl transferase complex like RbBP5 and WDR5. Though MLL is generally considered as the transcriptional activator, in this complex it was present along with transcriptional repressors like HDAC. More interestingly, except for MLL, there are no other Trithorax group proteins present in the complex.

Yokoyama et al. biochemically purified MLL along with the core components of SET1 like, complex. The proteins present in the complex are RbBP5, ASH2L, WDR5 and DPY-30. The SET domain of MLL is not intrinsically active for the H3K4 methyl transferase activity. These proteins are required for the substrate recognition and enzymatic activity of the MLL SET domain and thus for the transcriptional activation of MLL target genes. Additionally, this protein also contains HCF1 and Menin. In contrast to the super-complex identified by Nakamura, this 1MDa complex is much smaller in size and is made exclusively of transcriptional activators [59].

Biological Functions of MLL

MLL belongs to the Trithorax group (trx-G) proteins that are positive transcriptional regulators, which antagonize the function of Polycomb (PcG) mediated transcriptional repressors [63]. The trx and Pc-G proteins, which maintain the identity of cells by keeping the target genes active or silent respectively are key components of the epigenetic mechanisms that dictate cell fate determination [64]. The *TRX* mutants in *Drosophila* show anterior or posterior segmental transformations[65]. Yu et al. generated

Mll hypomorph mice by inserting promoter less *lacZ* in the third exon of *Mll*. The homozygous deletion of *Mll* (*Mll*-/-) is embryonically lethal at around day 10.5. *Hox* genes are expressed in *Mll*-/- mice until day 9.5, but their expression was not maintained any further, showing that Mll is required for the maintenance rather than initiation of *Hox* gene expression. Mice heterozygous for *Mll* (*Mll*+/-) were smaller in size, anemic and had defects in the axial skeleton, disordered segmental identities of cervical, thoracic and lumbar regions; highlighting the importance of MLL in development and hematopoiesis [8]. Mice with homozygous deletion of the SET domain of Mll also have altered body segments, decreased H3K4 mono methylation and decreased *Hox* gene expression [66]. The altered segment phenotypes are features of homeotic transformation, mirroring the phenotype of *Trx* mutations in *Drosophila*. Furthermore in mammals, *Mll* is required for normal hematopoiesis, skeletal and neural development. *Mll*-/- yolk sac and fetal liver have reduced cellularity and the hematopoietic progenitors from fetal livers show decreased expansion in methyl-cellulose colony formation assays [6]. This showed that Mll is critically required for fetal hematopoiesis, specifically required for the fetal HSC self-renewal. Using a conditional knock out of *Mll*, McMahon et al. showed that Mll is also required for self-renewal of HSCs postnatally [67].

Transcriptional Targets of MLL

Genome wide MLL occupancy and *in vivo* mouse studies and have been undertaken to identify the targets of MLL. Several studies have shown that MLL binds to a specific set of target gene promoters and through its intrinsic histone methyl transferase activity; and associated acetyl transferase activities activates the transcription of these

target genes. However, genome wide ChIP on chip for human MLL performed by Guenther et al. showed that 90% of RNA PolII bound promoters of actively transcribing genes were also bound by MLL. Furthermore 92% of MLL-bound promoters showed enrichment for H3K4Me3 suggesting a global role for the histone methyl transferase function of MLL in transcription. Interestingly, the findings from this study showed MLL occupancy and H3K4 trimethylation at *HOX* genes like *HOXA7, HOXA9, HOXA10, HOXA11 and HOXA13* was not just limited to the transcription start site and 5' region of the gene. Instead the chromatin was extensively occupied by MLL both upstream and downstream of the genes, and this binding correlated with active transcription of the genes. This was also shown for another MLL target gene *Meis1* [68]. Milne et al. showed that the C-terminus of MLL interacts with the Ser5 phosphorylated RNA polII CTD suggesting a global role for MLL in transcription. However, they also showed that MLL occupied only a subset of actively transcribing genes. ChIP performed for Mll and Ser5 phosphorylated CTD in MEFs and showed that they co-localized at Mll target genes like *Hoxa9* and *Hoxc8* [69].

In striking contrast to Guenther et al.'s finding, a genome wide ChIP on chip performed by Wang et al. in *Mll* +/+ and *Mll*-/- MEF showed that only 5% of promoters showed a Mll-dependency for H3K4 trimethylation. In agreement with a majority of studies in the MLL field, this study also identified a sub-set of *Hox* genes like *Hoxc8, Hoxc9* and *Hoxc6* as *Hoxc5*, *Hoxb4*, *Hoxb6*, *Hoxb10*, *Hoxa*6, *Hoxa7*, *and Hoxa9* as key MLL target genes. Moreover, MLL and MLL4 (also called MLL2) show some level of redundancy in regulating the transcription of their targets. Additionally, this study also identified that a major fraction of the MLL target genes are transcription factors and also

identified a couple of novel targets in the Wnt signaling pathway which in turn regulate the transcription of *Hox* genes [70].

Though the genome wide studies for the target genes of MLL show some contradictions, the *in vivo* studies from *Mll* mutant or knockout mice studies have clearly demonstrated *Hox*-specific phenotypes. *Hox* genes code for evolutionarily conserved homeo-domain containing transcription factors. There are 39 *Hox g*enes in mammals, arranged in four clusters (A to D) present in four different chromosomes. They are required for positional identity of body segments along the anterio-posterior axis of the body [71]. Hence they have a pivotal role in fetal development and embryogenesis. *Hoxa5, Hoxa7, Hoxa9, Hoxa10, Hoxb3, Hoxb6* and *Hoxc8* are the *Hox genes* that have been shown to be critical for hematopoietic stem cell (HSC) self-renewal [72]. Though many *Hox* genes are required for normal hematopoiesis, a few *Hox* genes like *Hoxa9, Hoxa7* and *Hoxa10* are over expressed in *MLL*-rearranged leukemia and are necessary for the leukemic transformation of hematopoietic progenitors.

Meis1 is another MLL target gene that encodes for a TALE (Three Amino acid Loop Extension) sub-type of homeo-box containing transcription factor [68,69]. Unlike *Hox* genes, the *Meis* family genes are non-clustered. *Meis1* collaborates with *Hox9-13* to bind DNA [14]. It is over expressed along with *Hoxa9* in *MLL*-rearranged leukemia and co-operates with *Hoxa9* in leukemic transformation of hematopoietic progenitors [15].

RUNX1 (*AML1*) is another gene that is frequently rearranged in AML leukemia. Like MLL, the RUNX1 protein is also required for normal hematopoiesis. Recently Huang et al., identified that an MLL interaction with the N-terminus of RUNX1 protects

RUNX1 from ubiquitin mediated proteosomal degradation. Moreover, MLL is required for maintenance of H3K4 trimethylation at the promoter of RUNX1 target genes like *PU.1*[73]. Besides *Hox* genes, cell cycle regulators like Myc and E2F1 target genes are also targets of MLL. Using a *Taspase1*-/- mice, Hsieh et al. demonstrated that proteolysis of MLL is required for transcriptional regulation of its targets. *Taspase1-/-* mice showed homeotic phenotype similar to *Mll* mutant mice, in addition, the *Taspase1* null mice were smaller in size. Furthermore, this study identified E2F target genes as transcriptional targets of MLL, highlighting a role for MLL in cell cycle regulation [74]. MLL is also a transcriptional regulator of cell cycle inhibitors like *Cdkn1a* and *Cdkn1b* [10,11].

Chromatin Accessibility and Histone Modifications

The human genome is made of more than 3 billion base pairs. To fit into a cell nucleus that is about 2µm in diameter, DNA has to be compacted at least 10,000 fold. This is made possible by wrapping of DNA around histones[75]. An histone octamer made from four different histones: H3, H4, H2A and H2B. About 146 bp of DNA are wrapped around a histone octamer forming a mononucleosome. Each mono nucleosome is connected to another by a linker histone, this arrangement or packaging fashion is referred to as beads on a string model which further coil helically to form a higher order solenoid in which the diameter of the chromatin fiber is 30 nm [76,77]. Though histones are very useful in compaction of DNA, they limit the accessibility of DNA for various DNA-templated cellular processes like transcription, DNA replication, DNA damage-response repair, recombination etc [78,79]. Chromatin associated proteins are key modulators of DNA accessibility. Post translational modifications of histone tails, exchange of core histones

for histone variants, sliding of nucleosomes by nucleosome remodeling complexes are the mechanisms that regulate the accessibility of DNA to regulatory proteins.

A portion of the amino terminal region of core histone proteins protrudes from the nucleosome, and is accessible to histone modifying enzymes. Post translational histone modifications act as docking sites for proteins that recognize the histone code and recruit other effector proteins [80]. The enzymes that post translationally modify the histones are referred to as writers. The proteins that recognize the modifications are referred to as readers. The proteins that remove the modifications are referred to as erasers. The basic amino acids like lysine and arginine in histones can be methylated. Additionally, lysines can be acetylated or ubiquitinated. The serines and threonines can be phosphorylated.

The effect of histone modification on transcription is very well studied. The transcriptional outcome depends on the amino acid modified, the histone tail the amino acid is present in (H3, H4, H2A and H2B), and the nature of the modification [81]. Acetylation of lysines in histones is considered as a transcriptional activation mark [82]. HATs like p300, CBP, and MOF are the enzymes that acetylate the lysines on histones, and the mark is removed by histone deacetylases (HDACs Class I, II and III) [83,84]. Acetylated lysines are recognized by bromo-domain containing proteins[85].

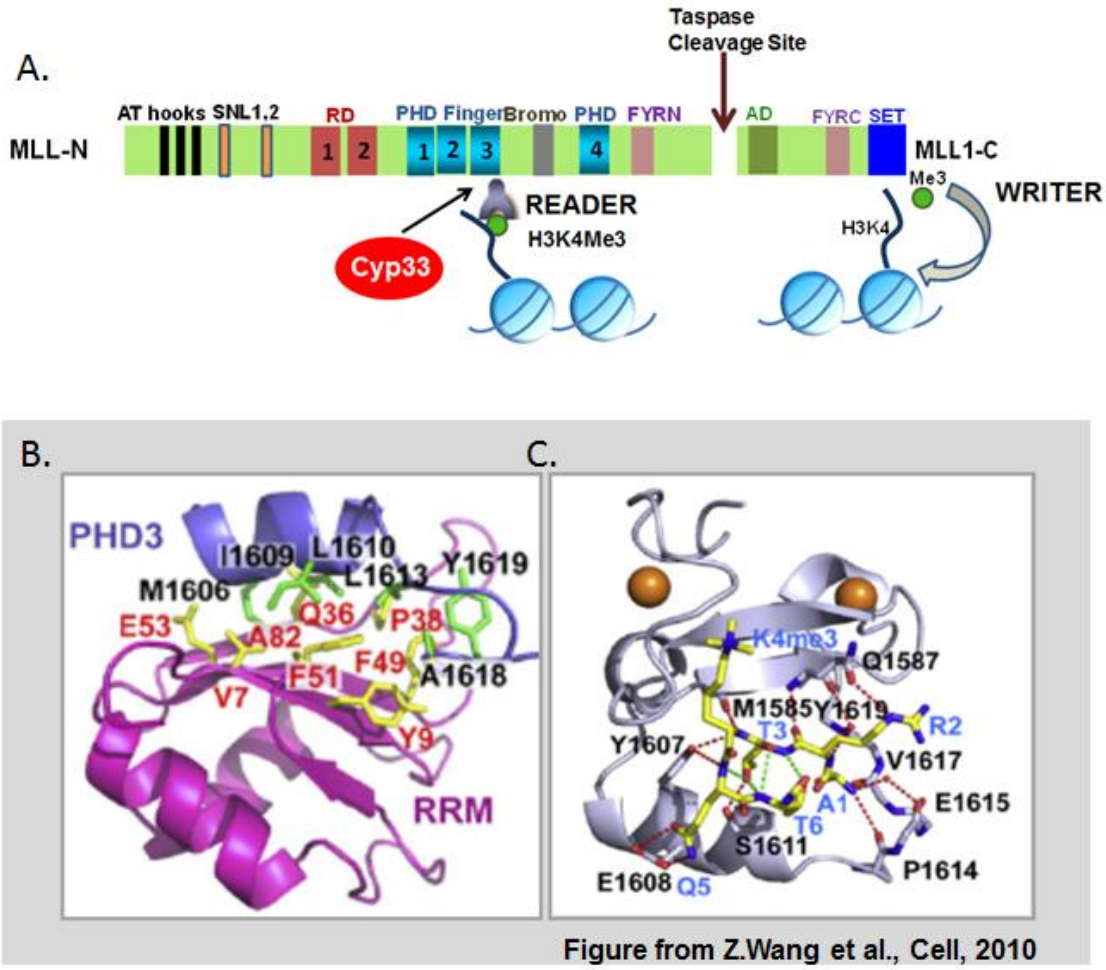
Histone methyl transferases; mono-, di- or trimethylate lysines on histones and have a SET (an acronym framed from the histone methyl transferases in *Drosophila* Su(Var)39, Enhancer of Zeste and Trithorax) domain as the catalytic domain. Some lysine methylation like H3K4Me3 is correlated with active transcription, whereas H3K9Me3 and H3K27Me3 are correlated with gene repression. The chromodomain,

Tudor domain, WD-40 repeat and PHD fingers recognize methylated lysines [85]. The Set1 enzyme found in yeast mono-, di-, and tri-methylates histone H3K4. SET1A and SET1B MLL, MLL2, MLL3, and MLL4 are the H3K4 specific histone methyl transferases in mammals[4]. The mammalian histone demethylase that demethylates mono- and di-methylated H3K4 is LSD1.The mammalian demethylases JARID1A, JARID1B, JARID1C and JARID1D demethylates mono-, di-, and tri-methyl histone H3K4 [86].

Cyp33 and H3K4Me3- The Known Ligands of MLL PHD3

The known ligands of MLL PHD3 are Cyp33 and H3K4Me3 (Figure 3A). Structural and NMR studies of MLL PHD3 binding to Cyp33 or H3K4Me3 were characterized by three different groups [26,29,30,87]. The *in vitro* binding studies from our collaborators Sangho Park and John Bushweller and from another group, have shown that Cyp33 has a higher binding affinity for MLL PHD3 than H3K4Me3 [26,30,87]. Isothermal calorimetry studies show that the MLL PHD3 binds to Cyp33 RRM with a kDa of 14.7 μM while the kDa for H3K4Me3 is 30.9 μM. Cyp33 and H3K4Me3 bind to two distinct surfaces of the MLL PHD3 but they are very close to each other. The effect of RRM of the Cyp33 on MLL PHD3-H3K4Me3 complex formation and vice versa shows that the interactions are mutually inhibitory [26]. The hydrophobic residues present on the α1 and α2 helices and on the loop 4 of the MLL PHD3, projects into the groove of β sheet of the Cyp33 RRM. This interaction is primarily mediated by hydrophobic residues V1617, Y1619, M1606, I1609 and L1610 in the MLL PHD3 and V7, Q36, F51, D53 and A82 in the Cyp33 RRM as shown in Figure 3B.

Figure 3: Ligands of MLL PHD3.



A, Schematic of MLL PHD3 binding to Cyp33 or H3K4Me3; **B,** A ribbon representation of the interaction of MLL PHD3 (blue) and Cyp33 RRM (magenta). The residues that are involved in the interaction in MLL PHD3 are in green and in Cyp33 RRM are in yellow. **C,** Crystal structure showing the hydrogen bonds between MLL PHD3 (blue) and bound H3K4Me3 peptide (yellow). This figure is from Z.Wang et al. Cell, 2010.

The high affinity interaction of the MLL PHD3 with the RRM of Cyp33 is inhibited by linking the MLL PHD3 to bromo-domain of MLL. However, this inhibition is overcome by adding the PPIase domain of Cyp33. The PPIase activity of Cyp33 targets the P1629 present in the linker region between the MLL PHD3 and the bromo-domain resulting in a conformational change that allows of the binding of MLL PHD3 to the RRM of Cyp33 [29]. The above mentioned finding is from *in vitro* binding studies. The importance of isomerization of P1629 on Cyp33 binding is yet to be demonstrated *in a* cellular system.

Similar to other PHD fingers that bind H3K4Me3, the MLL PHD3 also forms an aromatic cage that act as a binding module for the trimethyl ammonium group of K4Me3. The interaction is primarily mediated by Y1576, Y 1581, M1585 and W 1594 in the MLL PHD3. The binding of the MLL PHD3 to H3K4Me3 results in a conformational change in the loop segment between Y15756 to M1585A resulting in a conformational change that burries the methyl side chain of H3K4 in the aromatic cage. Moreover Wang et al. have shown that the bromo domain of MLL also contributes to the recognition of H3K4Me3 [29]. The MLL PHD3 interacts with the extended helix of the bromo-domain of MLL. In the absence of the interaction between MLL PHD3 and bromo domain only a partial aromatic cage is formed. The binding of the MLL PHD3 to H3K4Me3 contributes to the transcriptional activation of MLL target genes however the biological significance of this binding has yet to be determined.

Cyp33 and its Functions

Cyp33 is a 33 kDa, RNA binding cyclophilin family protein. It was first identified in 1996 from human T-cells [25]. It contains an RNA recognition motif (RRM) and a

Cyclophilin domain. The Cyclophilins are cyclosporin A binding proteins [88]. To date there are at least 16 cyclophilins identified in humans. All the cyclophilins have the prolyl isomerase activity. The proline in the peptide bond exist mostly in cis conformation, this is the restricted conformation. The peptidyl prolyl isomerase catalyzes the cis-trans prolyl isomerization of the prolyl bond and thus assisting in the proper folding of the protein [89]. Due to its peptidyl prolyl isomerase activity Cyp33 is also called PPIE or Cyclophilin E. The RRM of Cyp33 binds to poly-A or poly-U containing RNA. Cyp33 is a component of spliceosome and is involved in pre-mRNA splicing [24]. Besides binding RNA, the RRM of Cyp33 binds to proteins. Previous studies from our lab have shown that the RRM of human Cyp33 directly binds to the 3rd PHD finger of human MLL, and the interaction is not mediated by RNA binding. Moreover, the interaction between the Cyp33 and MLL PHD3 disrupts the interaction between Cyp33 and RNA [87]. Over expression of Cyp33 resulted in increased recruitment of transcriptional co-repressors like HDAC1 and Bmi-1 to the repression domain of MLL resulting in repression of MLL target genes [23,27]. More importantly this interaction is evolutionarily conserved and occurs in *Drosophila*. Dcyp33 binds to the third PHD finger of Trithorax, a *Drosophila* homologue of human MLL. Also, over expression of Dcyp33 in *Drosophila* SL1 cells resulted in down-regulation of the Trithorax target genes, highlighting the evolutionarily conserved role of Cyp33 in transcriptional regulation [57].

Figure 4: Illustration of Protein Domains of Cyp33.

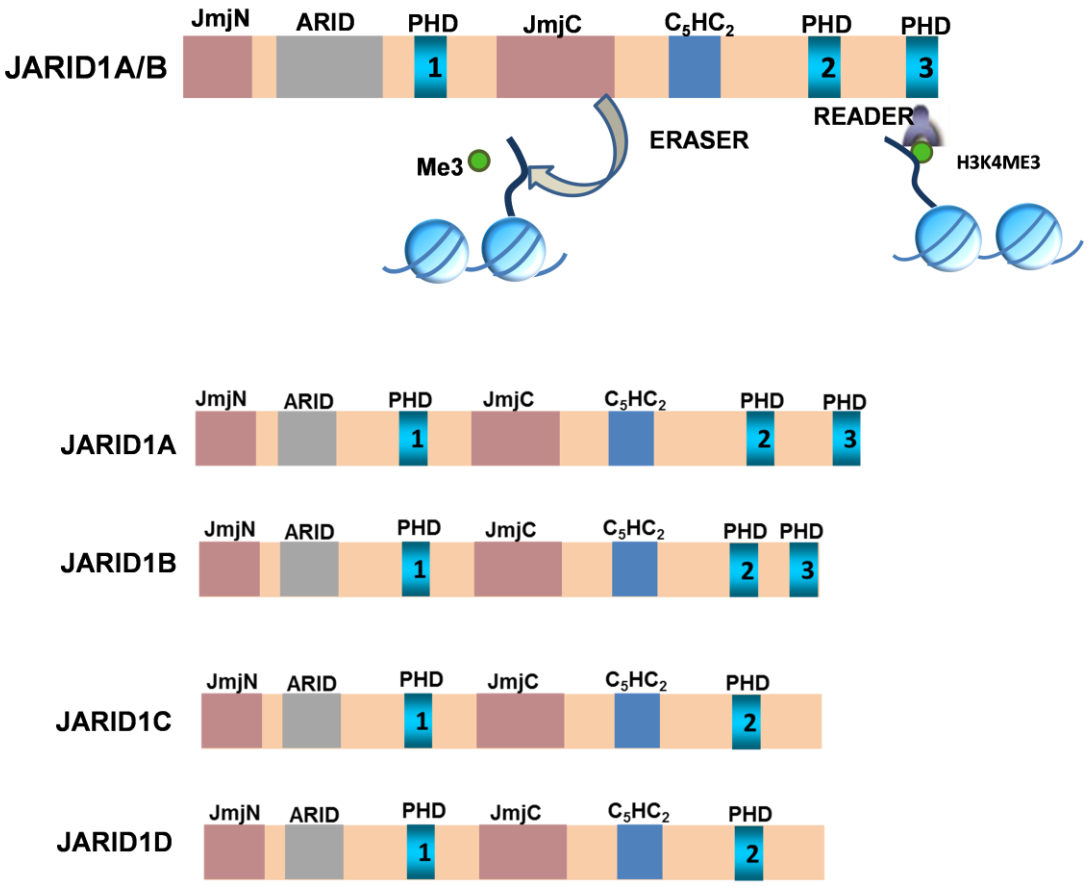
The N-terminal RRM domain of Cyp33 can bind to RNA and also mediates the interaction of MLL PHD3 and Cyp33. The Cyclophilin domain of Cyp33 contains prolyl isomerase activity. The region between the RRM and the Cyclophilin domain partially conserved and is known as the spacer region. The function of this region is unknown.

Furthermore, the PPIase activity is required for Cyp33 mediated transcriptional repression of MLL target genes [26,29]. Though Cyp33 was identified as a prolyl isomerase in 1996, until recently the targets of the Cyp33 PPIase activity were not known. Wang et al., recently demonstrated that the proline1629 present in the linker region between PHD3 and the bromo-domain of MLL is a target of the Cyp33 PPIase activity. The cis-trans isomerization of proline 1629 resulted in drastic conformational change allowing MLL PHD3 to bind to RRM of Cyp33, showing the requirement of PPIase activity for the transcriptional repression of target genes [29].

JARID1 Histone Demethylases and their Functions

Until the first H3K4 demethylase LSD1 was identified in 2004, the main mechanism to erase methyl marks was thought to be through passive dilution through successive cell division. LSD1 is a dual specificity demethylase that demethylates H3K4Me1 and H3K4Me2 and H3K9me1 and H3K9me2 [90,91]. In 2006 and 2007, a series of reports discovered the JARID1 (KDM5) family of mammalian demethylase that demethylates mono-, di-, and tri-methyl histone H3-K4. The JARID1 family also known as the KDM5 (Lysine demethylase5) family comprises of JARID1A, JARID1B, JARID1C and JARID1D. Like MLL, these proteins are also conserved throughout metazoan. The *Drosophila* homologue of JARID1 is called *Lid* (Little Imaginal Disc) [92]. All JARID1 family members have JmjN, ARID domain (AT Rich DNA Interacting Domain), PHD, zinc finger domains and a catalytically active JmjC domain that has the demethylase activity.

Figure 5: JARID1 family histone demethylases.



JARID1 family members, the histone demethylases remove trimethyl (Me3) mark from histone H3K4. The Jumonji-C (JmjC) domain contains the demethylase activity; the ARID domain is required for DNA binding. The PHD3 of JARID1A and JARID1B bind to H3K4Me3. Thus JARID1A and JARID1B are erasers with embedded readers for H3K4Me3 and are different from other two family members JARID1C and JARID1D that lack a PHD3.

The JmjN domain, ARID domains and zinc finger domains are all required for the demethylase activity of JmjC domain. The ARID domain binds to AT rich DNA sequences and thus imparts some specificity in binding to DNA. Of all the family members, JARID1A and JARID1B have similar domain architecture and are partially redundant in function [93]. Both JARID1A and JARID1B are "erasers" with embedded "readers". There are three PHD fingers in JARID1A and B. JARID1C and JARID1D lack PHD3. The 3rd PHD finger of JARID1A and JARID1B act as a ligand for H3K4Me3. The 2nd PHD finger of JARID1C specifically binds H3K9Me3 [35]. JARID1A is ubiquitously expressed, whereas JARID1B is expressed at low levels ubiquitously, except for the testis where there is higher level of expression. JAR1D1C and JARID1D are encoded by genes on the X and the Y chromosomes respectively. JARID1B was identified as an oncogene that is frequently over expressed in breast cancer. There is mounting evidence showing that JARID1B is over expressed in other human cancers like prostate, bladder, renal, and lung cancers, AML and CML. Though JARID1A and JARID1B are implicated as oncogenes, JARID1C and D are considered as potential tumor suppressors. There are inactivating mutations in JARID1C in kidney cancer and JARID1D is deleted in prostate cancers [86].

PHD Fingers of Other Chromatin-Associated Proteins

The PHD finger was first identified in the homeo domain containing protein HAT 3.1 in Arabidopsis and hence named as Plant Homeo Domain [94]. PHD fingers are about 60- 65 amino acids in length typically. However some atypical PHD fingers are about 80 amino acids long. More than 200 proteins in the nucleus contain PHD finger and a

majority of them are transcription regulators and chromatin associated proteins [95]. They have conserved regularly spaced cysteines, similar to that of the zinc fingers present in many transcription factors. Cys4–His-Cys3 is the amino acid sequence in a PHD finger that is required to co-ordinate two zinc atoms. Some atypical PHD fingers like that of RAG2 have Cys3-His2-Cys3 architecture to co-ordinate the zinc atoms [96]. Besides its role in DNA binding and mediator of protein–protein interactions, the biological function of PHD fingers have remained elusive since its discovery in 1993. The research over the past decade has highlighted the PHD fingers as readers or binding modules for modified and unmodified histones. In 2006, at least 6 different groups reported that the conserved aromatic residues present in PHD fingers of BPTF, NURF, ING2 and Yng1 forms an aromatic cage which serves as a binding pocket for trimethylated histone H3K4 [35,97,98]. Now there are about 15 different chromatin associated proteins that have PHD fingers that recognize H3K4Me3 (Table1) [35].

In addition to binding to methylated H3K4, the PHD fingers are now known to bind unmethylated H3K4, H3K36me3, H3K9Me3, H3K14 ac and H3R2 [35] . PHD fingers of BPTF, RAG2 and BHC30 recognize multiple but specific histone modifications. This combinatorial modification recognition by a PHD finger is plausible because the ligands (histones) may bind to different regions in the PHD finger and may influence one another's binding positively [98]. Also, multiple modifications are recognized by the same protein; for example, the bromo-domain of BPTF binds acetylated lysines and the PHD finger binds to H3K4Me3. In many of the chromatin associated proteins there are more than one PHD finger. The importance of the adjacent PHD fingers present in a same protein in recognizing the cognate histones is not yet well characterized. DPF3b PHD

fingers are the only known example wherein the PHD finger1 and 2 of DPF come together and form a binding pocket to recognize H3K14ac (Zeng et al. 2010) [99].

The PHD fingers are present in writer proteins that post translationally modify histones, eraser proteins that remove the histone modifications or may be present in a chromatin associated protein that recruits other transcriptional effector proteins. Hence the functional outcome of the PHD finger recognizing its cognate histone depends on the protein the PHD finger is present.

Mutations of PHD Fingers

Point mutations in the PHD finger region (N216S, V218I, and G221V) of ING family tumor suppressors are found in human cancers. These mutations result in about 10-40 fold decrease in the PHD finger binding to H3K4Me3 (Baker et al. 2009). Point mutations in the PHD finger region of RAG2 which affect the H3K4Me3 binding leads to decreased catalytic activity of the recombinase enzyme which is important for V(D)J rearrangement required for the development of B/T cells. An immunologic disorder, Omenn syndrome is a classic example where in the mutations in the PHD finger of RAG2 affect the H3K4Me3 binding are common. Especially there is higher incidence of W453R mutation in patients with Omenn's syndrome. This mutation specifically occurs in one of the three amino acids that contribute to the aromatic cage formation that disrupts the RAG2 PHD and H3K4Me3 interaction [100]. PHD fingers that bind H3K4Me3 can become a part of another protein by chromosomal translocations. Thus the fusion protein will have the acquired ability to bind H3K4Me3. The NUP98-JARID1A fusion occurs in rare cases of AML. NUP98 is a nucleoporin that shuttles between the nuclear

pore complex and regions of active transcription. The PHD finger 3 of the JARID1A histone demethylase or PHF23 that binds to H3K4Me3 can fuse to NUP98 resulting in a fusion protein. Both JARID1A and PHF23 PHD fingers can bind to H3K4Me3 explaining a common mechanism for leukemogenesis [101].

Table 1: PHD fingers of Chromatin-Associated Proteins that Bind H3K4Me3

Protein	Host Complex	Function of the complex (function of the protein)	Biological Outcome
BPTF	NURF	ATP-dependent chromatin remodeler	Nucleosome mobility Transcription activation
ING1	mSin3a/HDAC1	Histone deacetylase	Transcription repression
ING2	mSin3a/HDAC1	Histone deacetylase	Transcription repression
ING3	NuA4/Tip60	Histone acetyltransferase	Transcription activation
ING4	HBO1	Histone acetyltransferase	Transcription activation
ING5	MOZ/MORF HBO1	Histone acetyltransferase	Transcription activation
JARID1A (KDM5A)		(Histone demethylase)	Transcription repression
KDM7A		(Histone demethylase)	Transcription activation
KIAA1718		(Histone demethylase)	
MLL1	COMPASS-like	(Histone methyltransferase) Histone methyltransferase	Transcription activation
PHF2		(Histone demethylase)	Transcription activation
PHF8		(Histone demethylase)	Transcription activation
PHO23	Rpd3	Histone deacetylase	Transcription repression
PYGO1/2	PYGO1/2/BCL9	Transcription factor Wnt signaling	Transcription activation
RAG2	RAG1/2 V(D)J	Recombinase	Recombination
TAF3	TFIID	Transcription factor	Transcription activation
YNG1	NuA3	Histone acetyltransferase	Transcription activation
YNG2	NuA4	Histone acetyltransferase	Transcription activation

CHAPTER 3

MATERIALS AND METHODS

Cloning Techniques Used In This Study

Restriction Enzyme Digestion and Inactivation: Restriction endonuclease digestion of plasmid DNA was performed using the Fast Digest enzymes from Fermentas. 1-2µg of plasmid DNA was digested with 1U of the enzyme and 1µl of 10X Fast Digest buffer in a 10µl reaction for 10-15 minutes at 37°C. The enzyme was heat activated for 10 minutes; the temperature varies from 65-80°C, depending on the enzyme used.

Dephosphorylation of the vector: If the ends of the restriction digested vector DNA were compatible, the vector can re-ligate and form colonies upon transformation. To avoid this, after restriction digestion the vector fragment was dephosphorylated using 1µl of Fast Alkaline Phosphatase from Fermentas at 37°C for 10 minutes. The enzyme was heat inactivated at 80°C for 10 minutes.

Isolation of DNA restriction fragments from agarose gels: The digested DNA and an appropriate molecular weight marker were separated in 1% agarose gels in 1x TAE buffer (400mM Tris-acetate, 1mM EDTA). The gel was stained in ethidium bromide and destained in de-ionized water. The bands were visualized under UV and the DNA fragment of interest was identified based on the expected molecular weight. The DNA fragment of interest was excised from the gel. DNA was extracted from the gel slice using the Promega PCR clean up kit.

The extracted vector and insert DNA fragments were quantified by running on 1% agarose gel containing ethidium bromide and compared to the bands of known concentration and weight using a marker.

Ligation: The ligation was performed using T4DNA ligase from Fermentas. The gel purified vector and insert DNA fragments were mixed in 3:1 molar ratios in a 10µl reaction containing 1U of T4 DNA ligase from Fermentas and 1µl of 10X Ligase buffer containing ATP. Thermoligations were performed in a PCR machine overnight.

Transformation: 2µl of ligation reaction was electroporated into DH10B bacteria using a cell porator from GIBCO BRL according to the manufacturer's protocol. 500µl of LB media was added to the electroporated bacteria and incubated at 37°C with shaking at 225 rpm for 1 hour and then 100µl of the bacteria were plated on a LB ampicillin (100µg/ml) plate and incubated for 16-18 hours at 37°C.

Identification of the bacterial clone containing the ligated vector and insert plasmid: Individual colonies were picked and grown in 5ml of LB broth containing 100µg of ampicilin (or appropriate selection antibiotic resistance conferred by the plasmid). For larger constructs 50µg of ampicilin was used. The cultures were grown over night at 37°C in a shaking incubator at 225 rpm. The plasmid DNA was extracted from the bacteria using the Promega Plasmid Purification kit (mini preps). The plasmids containing the inserts in right orientation were identified by restriction digestion. To obtain larger quantity of DNA, Maxi prep was performed using Qiagen Maxi Prep Plasmid DNA purification kit.

Table 2. Expression Plasmids used in this study

Plasmid	Vector size (Kb)	Insert size (Kb)	Flanking Restriction sites	Protein length/ Amino acids	Source or Cloned by
pCep4	10.2	-----	-------	--------	Invitrogen
pCep4 FLAG Cyp33	10.2	3.0	KpnI…XhoI	Full length	Ute Osmers
pCep4 FLAG PPiase mutant	10.2	3.0	KpnI…XhoI	Full length (R191A, F196A)	Ute Osmers
pCXN2 FLAG MLL-N	5.2	6.5	EcoRI…XhoI	1-2158	James Hsieh
pcI FLAG MLL-N	5.7	6.5	XhoI	1 to 2665	James Hsieh
pGEX4t1 GST MLL PHD2	4.9	0.2	BamHI…XhoI	1480 to 1540	This study
pGEX4t1 GST MLL PHD3	4.9	0.32	EcoRI…XhoI	1531 to 1635	Ute Osmers
pGEX4t1 GST MLL PHD3 M1585A	4.9	0.32	EcoRI…XhoI	1531 to 1635 M1585A	This study
pGEX4t1 GST MLL PHD3 M1606D	4.9	0.32	EcoRI…XhoI	1531 to 1635 M1606D	This study
pGEX4t1 GST MLL4 PHD3 (ALL2)	4.9	0.32	EcoRI…XhoI	1335 to 1396	This study
MSCV neo	6.4	-----	--------	--------	Clontech
MSCV neo MLL-N wild type	6.4	8.0	XhoI	1 to 2665	Wei Wei Diaz lab
MSCV neo MLL-N M1585A	6.4	8.0	XhoI	1 to 2665	This study
MSCV neo MLL-N M1606D	6.4	8.0	XhoI	1 to 2665	This study
MSCV puro	6.3	-----	-------	--------	Clontech
MSCV puro FLAG MLL-ENL	6.3	4.8	EcoRI…XhoI	MLL 1-1394 ENL 372-560	Wei Wei Diaz lab
MSCV puro FLAG MLL-AF9	6.3	4.35	EcoRI…BglII	MLL 1-1362 AF9 481-568	Wei Wei Diaz Lab
pCMV Myc MLL C	3.8	3.75	XhoI	2720-3969	James Hsieh
Shcontrol pSIREN-RetroQ	6.4		BamH1…EcoRI	--------	Ralf Janknecht
Sh JARID1b pSIREN-RetroQ	6.4		BamH1…EcoRI	--------	Ralf Janknecht

Generation of MSCV neo MLL-N M1585A and M1606D

Step 1: Generation of MLL PHD3 mutants in pGEMT EZ

pCXN2 FLAG tagged MLL-N was kindly provided by Dr. James Hsieh from Memorial Sloan Kettering. A 2.3kb fragment of MLL-N which comprises the CXXC domain and PHD fingers was digested with AatII and NdeI and ligated into AatII and NdeI sites of pGEMT EZ (Promega). 20ng of this plasmid DNA was subjected to PCR based site directed mutagenesis using a forward primer containing the mutation and a reverse primer which anneals 200 bp downstream of the forward primer. The PCR products were digested with DpnI to get rid of the remaining input plasmid DNA and then transformed into DH10B bacteria. The transformants harboring the mutations were verified by DNA sequencing.

Table 3. Primers Used for Site Directed Mutagenesis

Primer Name	**Sequence 5`-3`**
MLL PHD3 M1585A F	GACTATGAGAGTAAGGCGATGCAATGTGGAAAG
MLL PHD3 M1606D F	GAGAATCTTTCAGATGAGGACTATGAGATTCTATCTAATC
MLL PHD3 R	GGAAGGTATACTCTCCTCTGTCTC

Step 2: Cloning of the full length MLL-N wild type and mutants in pBK(-)

Wei Wei in our lab cloned the 8.0kb wild type MLL-N into XhoI site of pBK (-) (Stratagene) from pcI MLL-N provided by Dr. Hsieh. The PHD3 mutants in pGEMT EZ were digested with AatII and NdeI, these 2.3kb fragments were swapped in for the wild type AatII …NdeI fragment in pBK(-) full length MLL-N.

Step 3: Cloning of MLL-N wild type and the mutants into MSCV neo

The 8.0kb MLL-N wild type or the mutants in pBK were digested with XhoI and ligated into the XhoI site of MSCV neo. The orientation of the insert was verified by restriction digestion and further confirmed by DNA sequencing.

Cell culture

The human cell line used in this study is HEK 293T. HEK 293T cells were grown in DMEM containing 10% FBS and 1% penicillin-streptomycin. The cells were maintained in culture by trypsinizing and splitting every 2 days. The mouse cell line used in this study is Mouse Embryonic Fibroblast (MEF). They are wt for MLL and grown in DMEM containing 10% FBS and 1% penicillin-streptomycin B-mercaptoethanol. Cells were trypsinized and re-plated every 2 days.

Transfection

HEK 293T cells were transfected using Lipofectamine-2000 (Invitrogen). Cells were trypsinized and plated in a 10cm dish at least 4 hours prior to transfection. The cells were approximately 40% confluent at the time of transfection. 15-20µg of plasmid DNA was transfected with 30µl of lipofectamine according to the manufacturer's protocol. DNA-lipofectamine complex was allowed to form at room temperature for 20 minutes in serum free media. The transfection mixture was added to the cells. The media was changed after 12 hours and the cells were harvested after 48 hours after transfection for experiments. If the transfected plasmids encode for drug resistance, then transfectants were selected using the specific drug (puromycin 1µg/ml for HEK 293T cells).

Retroviral packaging and concentration

$3x10^6$ ecotropic Phoenix virus packaging cells (Garry Nolan / Orbigen) were plated in a 10cm dish; five plates were plated per sample. The following day, 25µg of retro viral plasmid DNA per 10cm dish was transfected using calcium phosphate; to enhance the viral production 2.5µg of viral packaging plasmid was co-transfected. 14-16 hours post transfection 7ml of fresh DMEM containing 10% FBS and 1% Pen/Strep was added to each plate and the cells were incubated at 32°C with 5% CO_2. The Phoenix eco cells are adherent cells; the retroviral plasmid transfected cells release the viruses into the media. Hence the media containing the viruses were collected from the transfected cells and stored at 4°C. Fresh media was added to the cells, the cells were incubated at 32°C with 5% CO_2 to make more viruses. After 24 hours the media was collected again and combined with the media already collected, then spun down to remove any cell debris. The solution containing the retroviruses were concentrated about 7 times using Centricon Plus 70 filters (Millipore) by centrifugation in a swing bucket JS 7.5 rotor. Concentrated retroviruses were aliquoted, snap frozen, and stored at -80°C. The plasmids used for retroviruses were MSCV neo, MSCV neo MLL-N wt, MSCV neo M1585A, MSCV neo M1606D, and MSCV puro MLL-AF9 (Table 2).

Titration of Retrovirus

$8.5x10^4$ Rat 1a cells were plated in 6 well plates in DMEM containing 10% FBS and 1% Pen/Strep. The following day, Rat1a cells were infected with serial dilutions of the virus from $1x10^3$ to $1x10^6$ with 16µg/ml of polybrene in 1ml of DMEM containing 10% FBS. After 2 hours, the media containing polybrene was removed from the plates

and replaced with DMEM and 10% FBS. 24hours post-infection, the media was removed and fresh media containing 1mg/ml of G418 was added to select for the cells that have been transduced with the viruses. After the no virus control cells died (about 10 days), the colonies formed by virus infected cells were washed with PBS and stained with 1ml of a solution of 0.1gm methylene blue in 60 ml Methanol for 5 minutes. Plates were rinsed, dried, and then colonies were counted to determine the virus titer.

Retroviral Transduction of MEFs

About $0.5x10^6$ were plated on 6 well plates. After 3-4 hours the old media was removed and replaced with 1ml of media (DMEM 10% FBS, 1% Pen Strep) containing 750µl of concentrated retro virus, 4µg/ml of polybrene, 1 M HEPES was added. The plates were incubated at 32°C with 5% CO_2 for 4 hours. Then the media containing the virus was removed and replaced with fresh MEF media. Since the plasmid packaged into the virus encodes for neomycin resistance, the transductants were selected using G418 1mg/ml.

Isolation of Cells from Mouse Bone Marrow

One C57Bl/6 mouse, age 8-12 weeks, was sacrificed by CO_2 inhalation. Tibias and femurs were collected in a cell culture dish with RPMI 2% FCS. Excess tissue was removed from bones with sterile gauze. 8ml RPMI 2% FCS was added to a 10cm plate. Using a 25 gauge 1.5" needle and media in the dish, bone marrow was flushed into the plate until bones appear white. Cells were separated from the rest of the tissue debri by passing through a mesh. Cells were spun down at 1500 rpm for 5 minutes at 4°C, then resuspended in 5ml 1X cold sterile red cell lysis buffer (10X RBC lysis buffer: 83gm

NH_4Cl, 10gm $NaHCO_3$, 0.37gm EDTA in 1l H_20) and incubated on ice for 5 minutes. Cells were spun down at 1500 rpm for 5 minutes at 4°C, then resuspended in 10 ml PBS 2% FCS to create a single cell suspension and then proceeded with C-Kit positive selection to enrich for the hematopoietic progenitors population.

C-Kit Positive Selection of Mouse Bone Marrow Cells

To select for the C-Kit+ hematopoietic progenitor cells in the mouse bone marrow, I followed the Easy Sep Mouse CD117 (c-Kit) Positive Selection Kit (#18757 Stem Cell). Cells were resuspended at $100x10^6$ cells per ml in PBS 2% FBS, in a Falcon 5ml Polystyrene Round-Bottom Tube. CD117 antibody (with Fc blocker) was added to cells at 70µl/ml cells. Solution was mixed and incubated at room temperature for 15 minutes. Easy Sep PE selection cocktail was added at 70µl/ml cells. Solution was mixed and incubated at room temperature for 15 minutes. Magnetic nanoparticles were mixed by pipetting 5 times, and then added to the cell solution at 50µl/ml cells. Solution was mixed and incubated at room temperature for 10 minutes. Cell suspension was next brought up to 2.5mL with PBS 2% FBS. Cells were mixed by pipetting 2-3 times. Tube without cap was placed in chilled Easy Sep magnet and allowed to sit for 5 minutes. Without removing the tube from the magnet, the supernatant was carefully removed. The tube was removed from magnet and 2.5 mL PBS 2% FBS was added to cells, pipetting 2-3 times to mix. Magnet separation was repeated for a total of four 5 minutes separations. Cells were resuspended in 1.5mL PBS 2% FBS and counted. Cells were plated overnight in a round bottom 96 well dish at 1 million per 1mL of RPMI 1640 media with 10% FBS (heat inactivated at 55°C for 1hour) , 1% Pen-Strep, and supplemented with100ng/ml

SCF, 10ng/ml IL3, 10ng/ml IL6, 10ng/GM-CSF. Sterile H_20 was added to wells surrounding the cells to keep them from drying out.

Retroviral Transduction of C-Kit+ Hematopoietic Progenitors

The day after C-Kit positive selection, 200,000 cells C-Kit positive hematopoietic progenitors, 1.5ml of concentrated retro virus, 15.0µl 1M HEPES, 8µg of polybrene (Sigma) were mixed and the volume was made up to 2ml with RPMI 1640 media containing 10% FBS. Further, spinoculated at 3000 rpm in Beckman GS6KR centrifuge at 32°C for 4 hours. After spinoculation, without disturbing the cell pellet, the supernatant containing the polybrene and virus was carefully removed. The cells were then resuspended in RPMI 1640 media with 10% FBS (heat inactivated at 55°C for 1hour) , 1% Pen-Strep, and supplemented 100ng/ml SCF, 10ng/ml IL3, 10ng/ml IL6, 10ng/GM-CSF; and plated in 96 well plate. To ensure an efficient transduction of hematopoietic progenitors, a second spinoculation was performed the following day and cells were resuspeded in RPMI 1640 media with 10% FBS (heat inactivated at 55°C for 1hour) , 1% Pen-Strep, and supplemented 100ng/ml SCF, 10ng/ml IL3, 10ng/ml IL6, 10ng/GM-CSF and plated in 96 well plate.

Liquid Culture of Transduced Hematopoietic Progenitors

The day after retroviral transduction of hematopoietic progenitors, a portion of the cells (about 70,000 rest 30,000 was used for colony assay) were selected with G418 1mg/ml (MSCV neo) or puromycin (MSCV puro) 1µg/ml depending on the resistance marker encoded by the plasmid DNA used for retro virus production and cultured in RPMI 1640 media containing 10% FBS (heat inactivated at 55°C for 1hour), 1% Pen-

Strep, 100ng/ml SCF, 10ng/ml IL3, 10ng/ml IL6, 10ng/ml GM-CSF. Every 2 -3 days depending on their growth rate the media was removed and replaced with fresh media containing the indicated cytokines and selection drugs. After 4 days of selection cells transduced with MSCV neo, MLL-N wt, M1585A, or M1606D were harvested for RNA extraction.

Hematopoietic Progenitors Colony Formation Assay

The day after retroviral transduction of hematopoietic progenitors, 30,000 cells were added to 3ml of methyl cellulose supplemented with 100ng/ml SCF, 10ng/ml IL3, 10ng/ml IL6 (Stem cell technologies). Additionally, 10ng/ml GM-CSF and the selection drug G418 (1mg/ml) or puromycin (1μg/ml) was added and mixed very well by vortexing. About 1ml of the mixture was plated on to 35mm petri-dish. The platings were performed in duplicates. About 10,000 cells were plated in each plate for MSCV neo, MLL-N wt, M1585A or M1606D; about 1,000 cells were plated in each plate for MLL-AF9 transduced cells. The plates were placed in humid chamber and incubated at 37°C with 5% CO_2 and left undisturbed for a week. The hematopoietic progenitors should proliferate and form colonies on methyl cellulose. After a week, the number of colonies in each dish was counted. The transformation of hematopoietic progenitor was assessed by dissociating the colonies and re-plating the cells in methyl cellulose successively.

RNA Extraction, DNAseI treatment, cDNA synthesis and qRT-PCR

$5x10^6$ to$10x10^6$ cells were pelleted by centrifugation. The loosened cell pellets were lysed in 1ml of TRIzol™ (Invitrogen). 200μl of chloroform was added to the samples and mixed very well. The samples were incubated at room temperature for 2-3

minutes and then centrifuged at 12000 rpm for 12 minutes at 4°C. The upper layer (the aqueous phase) was carefully transferred to a fresh tube and the RNA was precipitated by adding an equal volume of isopropyl alcohol followed by centrifugation at 12000 rpm for 10 minutes. The pellet was washed once with 75% ethanol prepared in diethyl pyrocarbonate (DEPC) treated water. RNA pellet was air dried for 5-7 minutes and dissolved in 50µl of DEPC treated water by incubating at 55°C for 5 minutes. Concentration of the RNA at 260nm was measured using Nanodrop 2000c.

The RNA extracted by the above mentioned method may have some amount of DNA which will confound the RT-PCR results. Hence, the RNA preparation has to be subjected to DNAseI digestion. About 3µg of RNA was digested with DNAseI (Fermentas) according to the manufacturer's protocol. The DNAseI was inactivated by adding 50mM EDTA at 65°C for 10 minutes. Each DNAseI treated RNA sample was split into two halves. One half was used for cDNA synthesis with reverse transcriptase; the other was used as a control for DNAseI treatment wherein cDNA synthesis was performed without adding the reverse transcriptase (no RT control). cDNA will not be made in no RT control , hence any amplification detected in this sample will be from the remaining DNA present in the RNA sample. cDNA synthesis was performed using the high capacity cDNA synthesis kit from Applied Biosystems according to the manufacturer's protocol. cDNA samples were diluted 1:5 in nuclease free water and qRT-PCR was performed with GoTaq Sybr green from Promega. PCR for a specific primer set was performed in triplicates for each cDNA sample in a 96 well plate in ABI 3700 real time PCR machine. Sybr green is a fluorescent dye that intercalates DNA, and fluorescent signal at every cycle is recorded. Cycle threshold (Ct) value is the measure of

the amplicon signals that are detected at the linear phase of amplification and are above the background. Ct value is obtained by analyzing the raw data using the ABI PRISM 7300 software. The ct values (cycle threshold value) were obtained by performing the analysis using ABI Prism 7300 software.

Table 4. Human qRT-PCR Primers

Primer Name	**Sequence 5'-3'**
ACTB F	CGTACCACTGGCATCGTGAT
ACTB R	GTGTTGGCGTACAGGTCTTTG
GAPDH F 261	ATGGCAAATTCCATGGCACCGT
GAPDH R 417	ATGGTGGTGAAGACGCCAGT
B2M F	TGCTGTCTCCATGTTTGATGTATCT
B2M R	TCTCTGCTCCCCACCTCTAAGT
HPRT1 F	TGACACTGGCAAAACAATGCA
HPRT1 R	GGTCCTTTTCACCAGCAAGCT
PGK1 F	CTGTGGGGGTATTTGAATGG
PGK1 R	CTTCCAGGAGCTCCAAACTG
TBP F	TATAATCCCAAGCGGTTTGC
TBP R	GCTGGAAAACCCAACTTCTG
HOXA9 F	TCCCACGCTTGACACTCACACTTT
HOXA9 R	AGTTGGCTGCTGGGTTATTGGGAT
HOXC8 F	TGGAAACCTGAAGGAGATGTGGGT
HOXC8 R	AAACAGGGAAGGAGAGGAAGGCAT
CDKN1B F	CCGGCTAACTCTGAGGACAC
CDKN1B R	CTTCTGAGGCCAGGCTTCTT
MEIS1 F	CAGAAAAAGCAGTTGGCACA
MEIS1 R	TCATGCCCATTCCACTCATA
MYC F	CCTACCCTCTCAACGACAGC
MYC R	CTCTGACCTTTTGCCAGGAG
CYP33 F	TCCAGGCCAGTTTGGTCAGATGAT
CYP33 R	CCGGCTTGTTCCCAATCTTGATGT
MLL F	ATGTGGAAAGTGTGATCGCTGGGT
MLL R	TGCCGCTCAGTACAGTTCACACAA
KDM5A F	TGTCCTAAATGTGTCGCCGAGGAA
KDM5A R	TGTGGGAACCATATGGACTGGCAT
KDM5B F	AATTTGGCAGTGGCTTTCCTGTCC
KDM5B R	CTGCTCCATCACTGGCATGTTGTT

Table 5. Mouse qRT-PCR Primers

Primer Name	Sequence 5'-3'
m Actb F	GGCTGTATTCCCCTCCATCG
m Actb R	CCAGTTGGTAACAATGCCATGT
m Gapdh F	GTGAACGAGAAGGACTATAACCC
m Gapdh R	GGCTGTGTAGGAATGGACTG
m Hoxa9 F	TACTATGTGGACTCCTTCCTGC
m Hoxa9 R	TTCCACGACGCACCAAACAC
m Hoxc8 F	GACGCCTCCAAATTCTATGGC
m Hoxc8 R	AGACGAGTTCTGATTTAAGTGGCC
m Cdkn1b F	CGCCAGACGTAAACAGCTCCGAATTA
m Cdkn1b R	AGAGGCAGATGGTTTAAGAGTGCC
m Cdkn1a F	TGTCTGAGCGGCCTGAAGATT
m Cdkn1a R	AGACCAATCTGCGCTTGGAGTGAT
m Meis 1 F	CTCCTCTGCACTCGCATCAG
m Meis 1 R	CCATGGCGTTGGTATGAGCT
m Myc F	CCCACCACCAGCAGCGACTC
m Myc R	GCGGAGGTTTGTGGCCT
m Pu.1 F	TGGAACAGATGCACGTCCTCGATA
m Pu.1 R	TCCAAGCCATCAGCTTCTCCATCA

Fold change in gene expression was calculated by normalizing the average ct value of gene of interest for a given sample to the ct value for a house keeping gene for the same sample. Relative fold change in gene expression was calculated by setting the control to one and comparing the change in expression of other samples with respect to the control.

Chromatin immunoprecipitation (ChIP) followed by quantitative PCR

$10x10^6$ cells in 10ml of media were cross linked with formaldehyde (1% final concentration) for 15 minutes at room temperature. The cross linking was stopped by adding 1ml of 1.25 M glycine. The cells were washed with cold PBS twice and pelleted at 2000 rpm. To the cell pellet 500µl cell-lysis buffer (3mM MgCl2, 10mM NaCl, 10mM Tris (pH 7.4) and 0.1% NP40) containing protease inhibitor (cocktail from Roche) was

added and incubated on ice for 15 minutes with intermitten vortexing every 5 minutes. The samples were centrifuged for 5 minutes at 2000 rpm at 4°C. The supernatant was removed and to the pellet 500µl of nuclear lysis buffer (1% SDS, 10mM EDTA, 50mM Tris (pH 8.0)) containing protease inhibitor was added and sonicated on ice to shear the DNA to an average length between 200 bp to 1000 bp. The sonication was performed using a Cole Parmer, High Intensity Ultrasonic Processor/Sonicator, 50 watt model with a 2mm tip set to a power of 4 or 5. 293T cells were sonicated at setting 4, 6 pulses, 5 times. MEFs were sonicated at setting 5. The samples were centrifuged for 10 minutes at 13,000 rpm at 4°C. The debri was removed. The supernatant (nuclear lysate) was used for the immunoprecipitations. 50µl of the nuclear lysate was diluted 10 times with ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl (pH 8.0), 167 mM NaCl) containing protease inhibitor, to which indicated amount of primary antibody (Table 6) and 20µl of Protein-A magnetic beads (Millipore) were added and mixed very well and immunoprecipitations were performed on an overhead turner overnight at 4°C. 10µl of nuclear lysate from each sample was kept aside as input to normalize for the amount of DNA present in each sample. The magnetic beads were allowed to attach to the walls of the tube using a magnetic rack (Millipore) for 30 seconds. The supernatant was discarded. The magnetic beads protein A/ antibody/ chromatin complexes were washed in 500µl the following buffers for 3-5 minutes with rotation at 4°C. Low Salt Wash Buffer: (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl (pH 8.0), 150mM NaCl), High salt wash buffer: (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl (pH 8.0), 500mM NaCl), LiCl Wash buffer: (0.25M LiCl, 1% IGEPAL-CA630 or NP40, 1% Deoxycholic acid (sodium salt), 1mM

EDTA, 10 mM Tris (pH 8.0)) and TE Buffer: (10mM Tris-HCl (pH 8.0)**,** 1mM EDTA). Bound DNA was eluted from the magnetic beads using 100µl elution buffer (0.1M NaHCO3, 1% SDS) and1µl of Proteinase-K 10mg/ml, mixed very well and incubated at 62°C for 2 hours with rotation. The Proteinase-K was inactivated at 95°C for 10 minutes. The protein-free DNA was eluted from the magnetic beads. The DNA was recovered and purified using Mol Bio PCR clean up kit according to the manufacturer's protocol. Similar to the IP samples the inputs were also processed for elution and Proteinase-K digestion and purification of DNA. The eluted DNA was diluted 5 times in nuclease free water and qPCR was performed in triplicates for each sample in a 96 well plate in ABI 3700 real time PCR machine with the primers specific for the gene of interest (Table 7 and Table 8) using the GoTaq Sybr green from Promega. The fold recruitment of a specific chromatin associated protein was calculated by normalizing the amount of chromatin immunoprecipitated DNA to the amount of input DNA, and subtracting from the background (IgG pull down) levels. The fold enrichment for the histone modifications was calculated by normalizing first to the input levels and then to the amount of histone H3 levels present at the region of interest.

Table 6. Antibodies Used in ChIP

Antibody	**Manufacturer/Catalogue #**	**Amount used/ ChIP**
Anti histone H3	Millipore 05-499 Abcam ab1791	5µl
Anti histone H3K4Me3	Millipore 17-614	3µl
Anti acetyl histone H3	Millipore 06-599	5µl
Anti acetyl histone H3K27	Millipore 07-360	5µl
Anti JARID1A	Abcam ab70892	7.5µl
Anti JARID1B	Abcam ab50958	7.5µl
Rabbit IgG (negative control)	Millipore 12-370	5.0µl
Mouse IgG (negative control)	Millipore 12-371B	5.0µl

Table 7. Human ChIP Primers

Primer Name	**Sequence 5'-3'**
ACTB Promoter F	TGCACTGTGCGGCGAAGC
ACTB Promoter R	TCGAGCCATAAAAGGCAA
GAPDH Promoter F	TACTAGCGGTTTTACGGGCGCACGT
GAPDH Promoter R	TCGAACAGGAGGAGCAGAGAGCGA
B2M Promoter F	AGACTTCCCAAATTTTGCCATCCTA
B2M Promoter R	AAAGGCCTGAAATGTTAGTGTTGAGT
BGLOBIN Promoter F	AGGGAGGGCTGAGGGTTTGA
BGLOBIN Promoter R	CAGGGTGAGGTCTAAGTGATGACA
Histone H4 Promoter F	AAATGGTGGGATCACAGACG
Histone H4 Promoter R	CGAGCTTCTTGTTTCCGTGT
RPL10 Promoter F	ACCCGTCTTCGACAGGACT
RPL10 Promoter R	GGAACGGAAGACGAGGAGAACAG
HOXA9 Promoter F	GCCAACCAAACACAACAGTC
HOXA9 Promoter R	CAAATCGCATTGTCGCTCTA
HOXC8 Promoter F	GGCGAGAGGCACCTTCACATCTAAATTC
HOXC8 Promoter R	AGAGAGAGAGAGAGAGAGAGAGAGAGATGGC
CDKN1B (p27) F	CCGGCTAACTCTGAGGACAC
CDKN1B (p27) R	ATACGCCGAAAAGCAAGCTA
JARID1B Promoter F	ACTTCTTCAGGGCAGGAACTCTGA
JARID1B Promoter R	TCCGCTGTCGATGGAGTTTATCCT
HDAC1 Promoter F	CGGCTATAGGTGAGCCCAGGA
HDAC1 Promoter R	AACAGACTTTCCTCCGGGT

Table 8. Mouse ChIP Primers

Primer Name	**Sequence 5'-3'**
m Actb Promoter F	AGCAATAGCCGGAAAGCCAGATCC
m Actb Promoter R	ACCACTGGGGCTCGCCCTATG
m Gapdh Promoter F	AGAGAGGGAGGAGGGGAAATG
m Gapdh Promoter R	AACAGGGAGGAGCAGAGAGCAC
m Hoxa9 Ch pro can F	CACTGCCTGAGACCACAGAA
m Hoxa9 Ch pro can R	AGGGAGCCTGGAGTACTGGT
m Hoxc8Promoter F	AACCTGAGTTGACCTGATGGCTCA
m Hoxc8 Promoter R	CTAACTAATGGACAAGGCCAGCTT
m Cdkn1b Promoter F	CGGCCGTTTGGCTAGTTTGTTTGT
m Cdkn1b Promoter R	GGAGGCTGACGAAGAAGAAGATGA

Western Blot and Immunoprecipitation

Cells were washed in PBS and pelleted down by centrifugation. Further, cell lysis was performed using IPH lysis buffer (50mM Tris-HCl (pH 8.0), 150mM NaCl, and 5mM EDTA, 0.5% NP-40, 1X protease inhibitor (Roche) for 30 minutes on ice. The cellular debri were removed by centrifugation at 12000 rpm 10 minutes at 4°C. Protein concentration was estimated using BCA (Pierce) method. Depending on the abundance of the protein in the cell, anywhere from 500 to 1500µg of protein was made up to 1ml with 1X IPH lysis buffer. To reduce non-specific binding, the lysate was pre-cleared with ProteinA/G beads (Roche), depending on the isotype of primary antibody used for 1hour at 4°C. The pre-cleared lysate was used for immunoprecipitation with the primary antibody (at the concentration recommended by the manufacturer) along with ProteinA/G beads (Roche) over night at 4°C. The immunoprecipitated protein complexes were washed with lysis buffer for 4-5 times with 5 minutes intervals between each wash. After the final wash, 2X loading dye (4% SDS, 10% 2-mercaptoethanol, 20% glycerol, 0.004% bromophenol blue, 0.125 M Tris-HCl) was added to the beads and denatured at 100°C and loaded onto SDS-PAGE (percentage of PAGE was ascertained based on the molecular weight of the protein to be immune blotted) and the protein was electrophorectically separated using the 1X Tris-Glycine running buffer (25 mM Tris base, 190 mM glycine, 0.1% SDS pH 8.0). The separated proteins were electrophoretically transferred to a PVDF membrane with a pore size of 0.45µM (Pierce) using 1X pre-chilled transfer buffer (25 mM Tris base,190 mM glycine, 20% methanol) 100 volts for 1 hour. After the protein was transferred to the membrane, to minimize the non-specific binding of the antibody, the membrane was blocked in 5% non-fat skim milk

in 1X TBS (24.23 g Tris-HCl, 80.06 g NaCl made up to 1L with water, pH to 7.6) for 1 hour at room temperature. Then, primary antibody was diluted as per the dilutions in the table in 2.5% milk and added to the blot and left in the shaker over night at 4°C. The following day the blot was washed in TBST (0.5ml of Tween 20 added to 1L of TBS), 10 minutes, three times. Depending on the organism in which the primary antibody was raised, the corresponding secondary antibody (anti Rabbit or anti Mouse from Santa Cruz) was added for 1hour at room temperature. To get rid of the unbound antibody, and other non-specific binding the blot was washed three times with TBST and the blot was developed with Enhanced Chemiluminescence kit (ECL Super signal, Dura) from Pierce).

Table 9. Primary Antibodies Used in Western Blot

Antibody	**Manufacturer/Catalogue #**	**Epitopes recognized**	**Dilution**
Anti H3K4Me3	Abcam 12209	1-100	1:1000
Anti MLL-N	Dr.Thirman	1-371	1:1000
Anti MLL-N	Bethyl lab	1320-1380	1:2500
Anti MLL-C	Dr.James Hsieh	2829-2883	1: 10000
Anti PPIE	Novus Biologicals H00010450-A01	2-101	1:1000
Anti JARID1A	Abcam ab70892	1675-1722	1:1000
Anti JARID1B	Abcam ab50958	125-325	1:1000
Anti-FLAG M2	SIGMA F3165	DYKDDDDK	1:1000
Anti HA	Abcam 9110	YPAYDVPDYA	1:1000

Western Blot for MLL-N

$8x10^6$ cells (MEF or 293T) were washed in PBS and centrifuged at 2000 rpm for 5 minutes. The pelleted cells were then lysed in 300µl of cell lysis (3mM MgCl2, 10mM NaCl, 10mM Tris (pH 7.4) and 0.1% NP40) buffer for 15 minutes on ice with intermitted vortexing followed by centrifugation at 3000 rpm for 5 minutes at 4°C. To the pellet,

350µl of nuclear lysis buffer (1% SDS, 10mM EDTA, 50mM Tris (pH 8.0)) was added and chilled in ice for few minutes and then sonicated on ice at setting 5 for 3 pulses two times. The nuclear debri was removed by spinning at 12000 rpm for 10 minutes at 4°C. The protein present in the nuclear extract was quantified by BCA (Pierce) method; 100µg of the nuclear extract was mixed with 6X loading dye, denatured at 100°C for 5 minutes and loaded on 3-8 % tris acetate gel from Invitrogen. The proteins were electrophorectically separated using the 1X tris- acetate running buffer (Invitrogen) at 160 volts for 1hour and 30 minutes. The separated proteins were electrophoretically transferred to a nitrocellulose membrane (High bond-C) with pore size 0.45µM (Amersham) using 1X pre-chilled transfer buffer containing 10% methanol (Invitrogen) at 100 volts for 3 hours. High molecular weight protein ladder from Fermantas was used a guide for assessing the separation of proteins on the gel and for the efficient transfer of protein to the nitrocellulose membrane. Following the transfer, blocking, primary antibody, secondary antibody staining and developing the membrane with ECL was done similar to the immune blotting for other proteins.

Nucleosome pull down assay

HEK 293T cells were transfected with 20µg of MSCV neo vector control; FLAG tagged MLL-N wt, M1585A or M1696D with Lipofectamine-2000 (Invitrogen) according to the manufacturer's instruction. After 48 hours about 1 to $2x10^7$ cells were washed with PBS and resuspended in a ml of buffer A (10mM HEPES pH 7.9, 10mM KCl, 1.5mM MgCl2, 0.34M Sucrose, 10% glycerol, 1mM DTT) containing protease inhibitor cocktail (Roche). Triton X-100 was added to the final concentration of 0.1% and

the resuspended cells were homogenized in a dounce homogenizer (about 8-10 strokes) and then centrifuged at 4°C, for 5 minutes at 1300g. The nuclei in the pellet were washed in 1ml of buffer A and then lysed in buffer B (3mM EDTA, 0.2mM EGTA,1mM DTT) containing PI and the soluble and insoluble chromatin fraction was separated by centrifugation at 4°C for 5 minutes at 1700g. The insoluble chromatin was resuspended in 0.2ml of buffer M (10mM Tris pH 7.6, 10mM KCl, 1.5 mM CaCl2) prewarmed at 25°C. 1U of micrococcal nuclease (New England Biolabs) was added and the digestion was performed for 5 minutes and 30 seconds. The reaction was stopped by adding 1mM EGTA and then centrifuged at 4°C 1700g for 5 minutes. An aliquot of the supernatant was de-proteinized using 1µl of 10mg/ml ProteinaseK at 62°C for 2 hours, followed by heat inactivation of Proteinase K at 90°C for 10 minutes. DNA was purified by PCR clean up kit and checked for predominant release of mononucleosomes in ethidium bromide containing 1% agarose gel. The length of the DNA wound around a nucleosome is approximately 146 bp. The remainder of the supernatant was adjusted to 0.5ml with buffer C (15mM tris pH 7.6, 150mM NaCl, 1mM EDTA, 1mM DTT) containing PI and incubated over night at 4°C with FLAG beads. The immunoprecipitaites were washed with buffer C 4 times and eluted by boiling the beads in 2X SDS sample buffer for 5 minutes. Aliquots of input and immunoprecipitated eluates were subjected to immunoblot analysis in a 15% SDS–PAGE and immunoblotted for H3K4Me3 (Abcam ab12209) according to the manufacturer's instructions [102].

CHAPTER 4

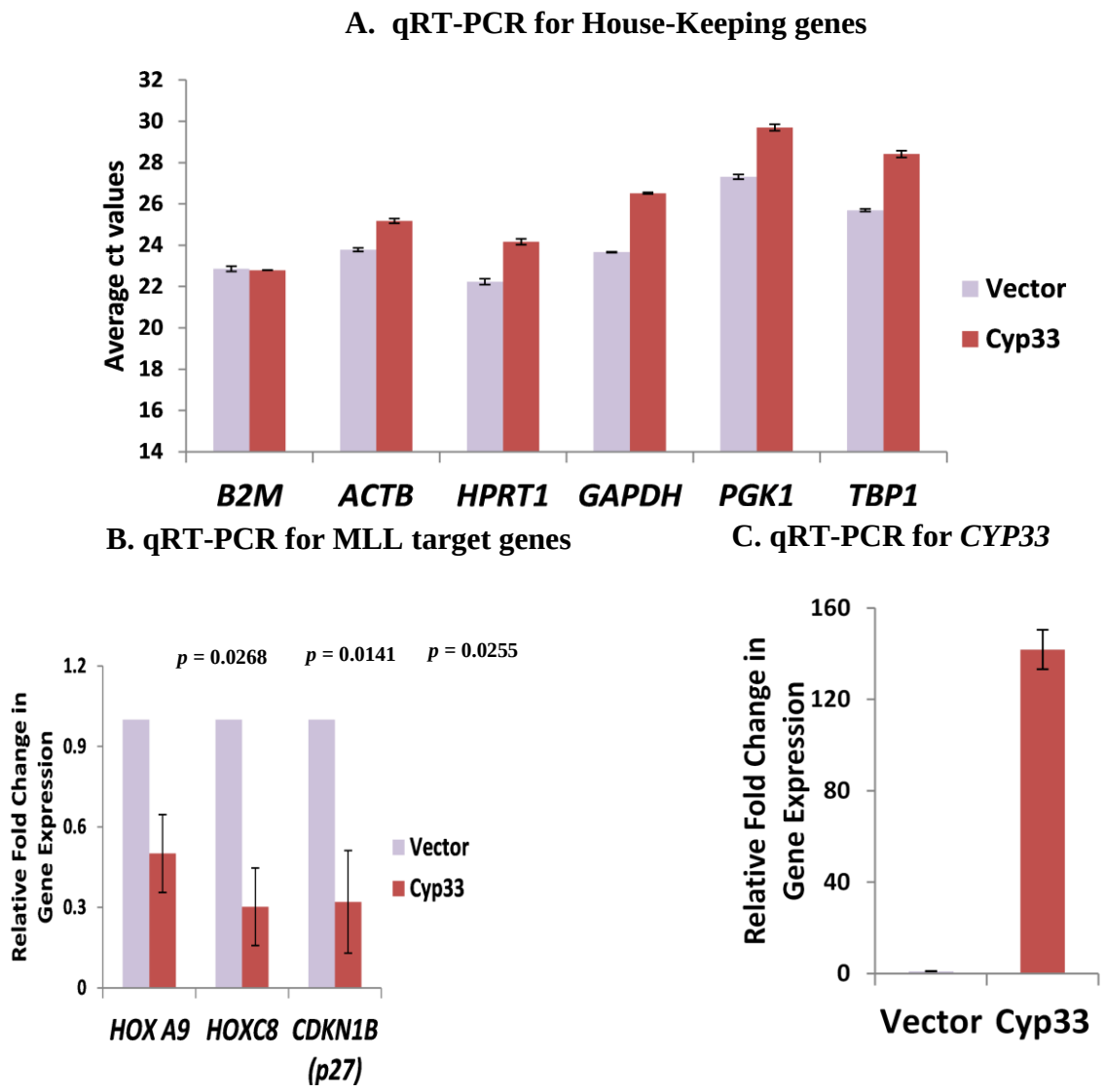
RESULTS

Aim1: Determine the effect of Cyp33 over expression on transcription of MLL target genes.

Previous studies from our laboratory have shown that the MLL PHD3 interacts with Cyp33 and over expression of Cyp33 results in transcriptional repression of MLL target genes. These previous studies used semi-quantitative RT-PCR to measure the effect of Cyp33 over expression on the transcription of MLL target genes [23,57]. We have confirmed those results using qRT-PCR, a more quantitative approach to measure gene expression. The effect of Cyp33 over-expression on a panel of house-keeping genes was tested by qRT-PCR (Figure 6A). Since the ct values of *β2- microglobulin* (*B2M)* mRNA levels remain unchanged between vector-transfected control and Cyp33 over-expressing cells, *B2M* is used as a reference gene. The effect of Cyp33 on the expression of MLL target genes was measured by qRT-PCR (Figure 6B). Cyp33 over expression in 293T cells resulted in a 2 fold decrease in expression of *HOXA9*, and 3.3 fold decrease in expression of *HOXC8* and *CDKN1B (p27)* all of which are known MLL target genes [11] Over-expression of *Cyp33* was verified using qRT-PCR for *CYP33* (Figure 6C).

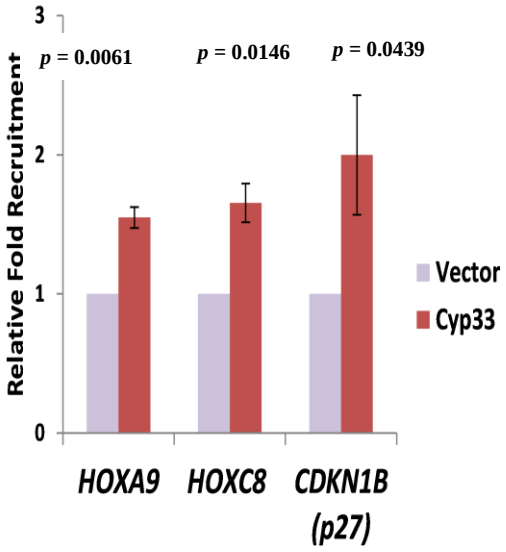
Nucleosome density and certain histone modifications are correlates of transcription. To understand the mechanism of Cyp33 mediated transcriptional repression of MLL target genes, ChIP for H3 and histone marks at the promoters of MLL target

Figure 6: Effect of Cyp33 over expression on transcription of MLL target genes



293T cells were transiently transfected with pCep4 (vector) or pCep4-Cyp33 plasmids and harvested after 48 hours for RNA extraction, followed by qRT-PCR: **A**, for house-keeping genes**; B**, for MLL target genes; **C**, for *Cyp33*. n=3, average of three independent experiments, error bars represent S.D. *p* values are from one- sample t- test.

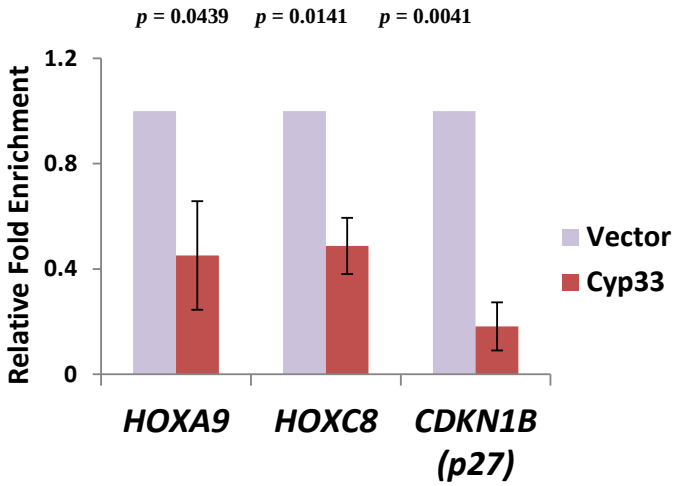
Figure 7: Effect of Cyp33 over expression on histone H3 density at MLL target gene promoters



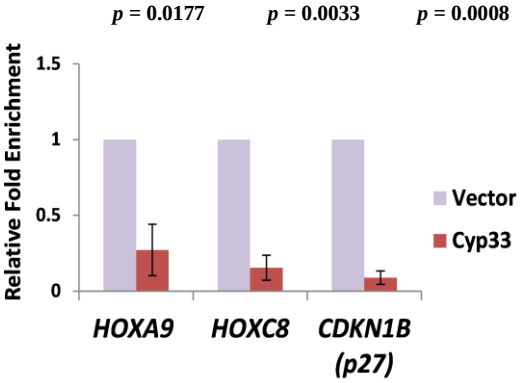
293T cells were transiently transfected with pCep4 (vector) or pCep4-Cyp33 plasmids and harvested after 48 hours for ChIP for histone H3; Rabbit or mouse IgG was used as a negative control for the pull-down. n=3, average of 3 independent experiments and error bar represents S.D. *p* values are from one sample t- test.

Figure 8: Effect of Cyp33 over expression on histone modifications at MLL target gene promoters

A. ChIP for H3Ac at MLL target gene promoters



B. ChIP for H3K4Me3 at MLL target gene promoters



293T cells were transiently transfected with pCep4 (vector) or pCep4-Cyp33 plasmids and harvested after 48 hours for ChIP for: **A**, H3Ac; **B**, H3K4Me3, at MLL target genes. The fold enrichment of histone H3 modifications was calculated by normalizing the amount of histone H3 modification to the amount of total histone H3 present at a given promoter. Rabbit or mouse IgG was used as a negative control for the pull-down. n=3, average of 3 independent experiments and error bar represents S.D. *p* values are from one sample t- test.

genes was performed. Over-expression of Cyp33 resulted in a statistically significant increase in histone H3 density at MLL target gene promoters which is consistent with the reduction in MLL target gene expression (Figure 7).

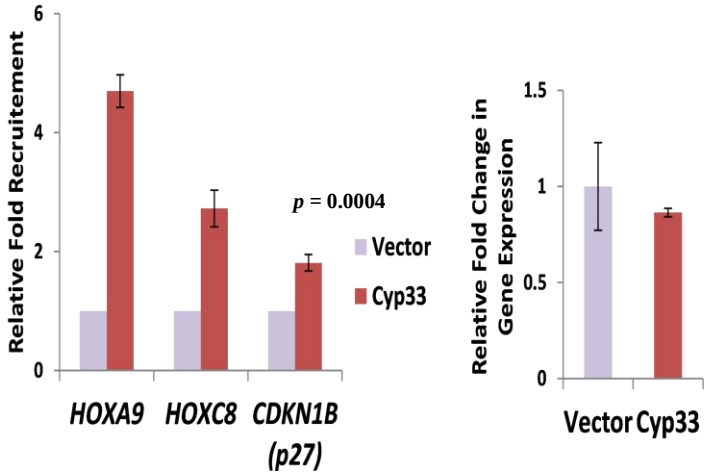
Studies from our laboratory have shown that Cyp33 over expression did not affect the histone marks like H3K27Me3 and H3K9Me3 associated with gene silencing. To test whether marks that are associated with active transcription change with Cyp33 over expression, a ChIP for transcription activation-associated histone marks like H3K4Me3 and H3Ac at the promoters of MLL target genes was performed. To determine the fold enrichment of a particular histone modification at a given promoter, the amount of a particular H3 modification is normalized to the amount of total histone H3 present at the given promoter. ChIP performed with a pan-acetyl H3 antibody showed that over expression of Cyp33 resulted in a decrease in H3Ac levels at the promoters of MLL target genes *HOXA9* (2 fold), *HOXC8* (2 fold) and *CDKN1B* (5 fold) (Figure 8A). ChIP for H3K4Me3 (Figure 8B) showed that upon Cyp33 over expression there is a marked decrease of H3K4Me3 at the promoters of MLL target genes *HOXA9* (3.7 fold), *HOXC8* (6.5), and *CDKN1B* (11.3 fold).

The Cyp33 mediated decrease in H3K4Me3 marks at MLL target gene promoters may be due to active demethylation of H3K4Me3 marks. JARID1 family proteins are histone demethylases that demethylate H3K4Me3 and H3K4Me2 [33,34]. To determine if Cyp33 mediated decrease in H3K4Me3 involved JARID1 histone demethylases, ChIP for JARID1A and JARID1B at MLL target gene promoters was performed.

Figure 9: Effect of Cyp33 on the recruitment of H3K4Me3 demethylases JARID1A and JARID1B

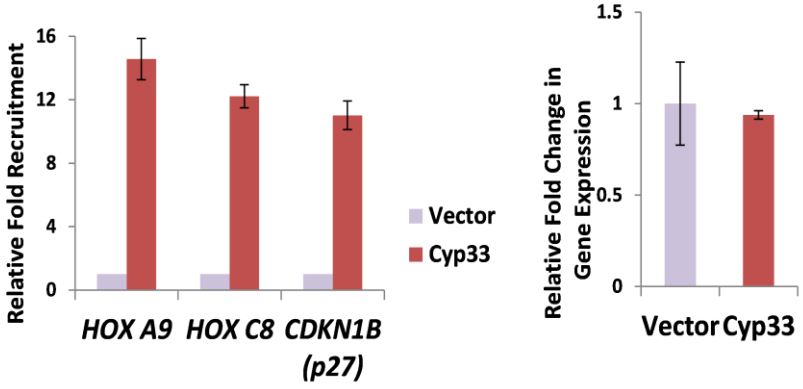
A. ChIP for JARID1A at MLL target gene promoters

B. qRT-PCR for *JARID1A*

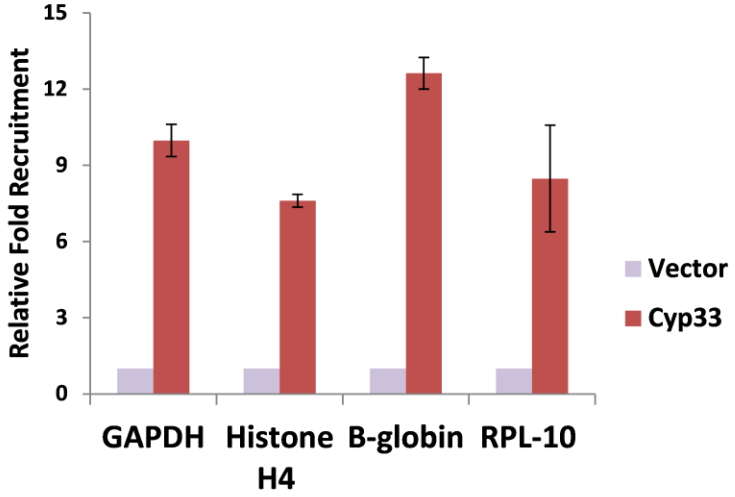


C. ChIP for JARID1B at MLL target gene promoters

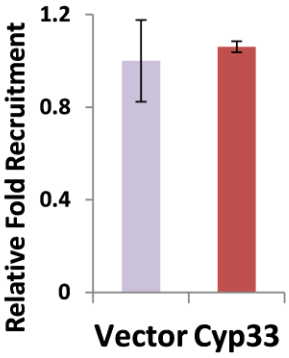
D. qRT-PCR for *JARID1B*



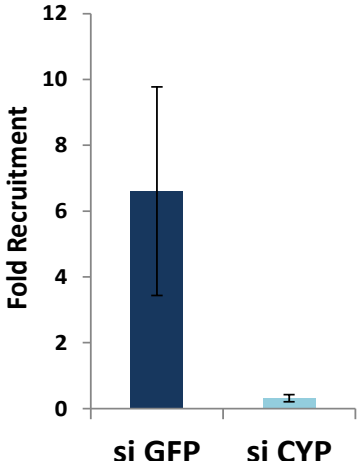
E. ChIP for JARID1B at House-keeping gene promoters



F. ChIP for JARID1B at *JARID1B* promoter



G. ChIP for JARID1B at *CDKN1B* promoter



H. qRT-PCR for *CYP33*

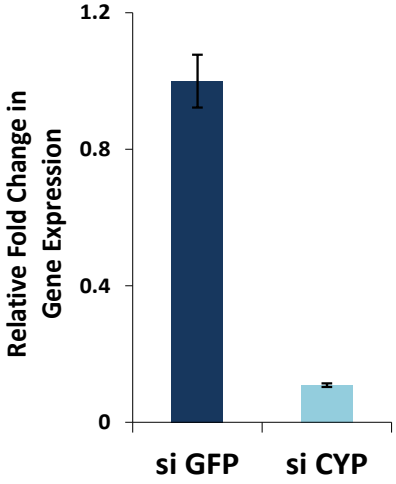


Figure A through F 293T cells were transiently transfected with pCep4 (vector) or pCep4-Cyp33 plasmids and harvested after 48 hours for RNA extraction or ChIP. **A,** ChIP for JARID1A at MLL target gene promoters; **B**, qRT-PCR for *JARID1A;* **C**, ChIP for JARID1B at MLL target gene promoters; **D,** qRT-PCR for *JARID1B;* **E,** ChIP for JARID1B at house-keeping gene promoters; **F,** ChIP for JARID1B at *JARID1B* promoter. n=3, average of 3 independent experiments, error bar represent S.D. Figure G and H MSA cells were transiently transfected with si RNA for GFP or CYP 33 after 48 hours for RNA extraction or for ChIP. **G,** ChIP for JARID1B at CDKN1B promoter; **H,** qRT-PCR for CYP33. n=1, the error bars are from triplicates from q-PCR.

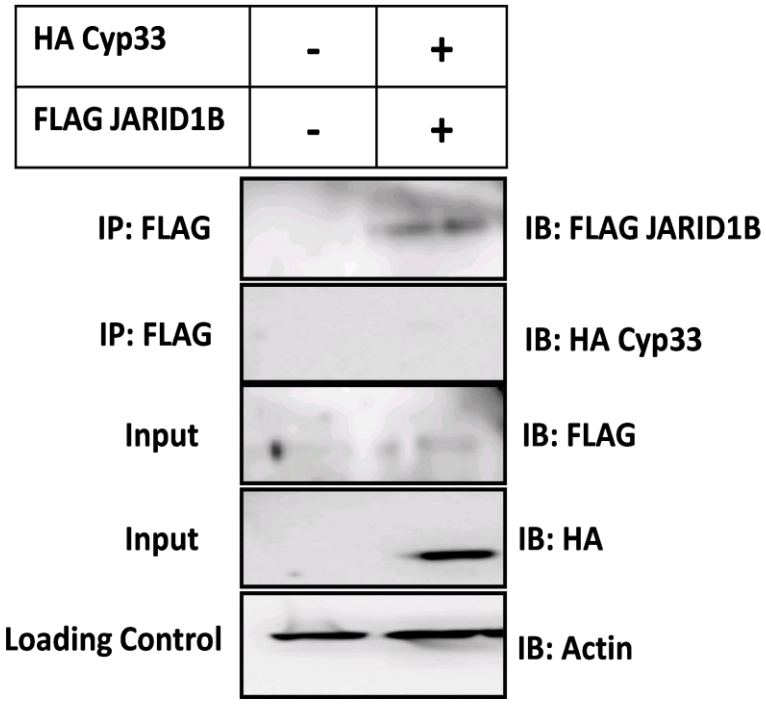
Upon Cyp33 over expression there is an increase in recruitment of JARID1A at the promoters of MLL target genes *HOXA9* (4.7 fold), *HOXC8* (2.7 fold), and *CDKN1B* (1.8 fold) (Figure 9A). Also Cyp33 over expression resulted in a pronounced increase on JARID1B recruitment at the promoters of MLL target genes *HOXA9* (14.5 fold), *HOXC8* (12.2 fold) and *CDKN1B* (11 fold) (Figure 9C). As shown by qRT-PCR Cyp33 over expression did not affect the overall expression of *JARID1A* and *JARID 1B* (Figure 9B and D).

To investigate if the Cyp33 mediated increase in recruitment of JARID1B occurred specifically at MLL target gene promoters, or also happens at other promoters, ChIP for JARID1B at the promoters of house-keeping genes like *GAPDH*, *Histone H4*, *B-globin* and *RPL10* was performed (Figure 9E)**.** Similar to MLL target genes, over expression of Cyp33 resulted in increased recruitment of JARID1B at the promoters of house-keeping genes. Since JARID1B expression was unaffected by Cyp33 over expression, there should not be increased recruitment of JARID1B at JARID1B promoter upon Cyp33 over expression. Hence ChIP for JARID1B at the *JARID1B* promoter was performed (Figure 9F). Cyp33 over expression did not increase the recruitment of JARID1B at JARID1B promoter. This shows that Cyp33 mediated increased recruitment of JARID1B does not occur globally at all promoters, it occurs at specific promoters, but not restricted to MLL-target promoters. The effect of Cyp33 over expression on JARID1B recruitment was confirmed by Cyp33 knock down experiment performed by Steven Poppen in the laboratory (Figure 9G & 9H). Knock down of Cyp33 resulted in decrease in JARID1B recruitment suggesting that JARID1B may be a downstream effector of Cyp33 mediated transcriptional regulation.

Similar to MLL; both JARID1A and JARID1B have a PHD3. Since Cyp33 RRM is a ligand for the MLL PHD3, we wanted to know if the same is true for JARID1A and JARID1B PHD3s. Hence a co-immunoprecipitation was performed in 293T cells (Figure 10). In contrast to MLL, JARID1B does not interact with Cyp33.

Figure 10: Verification of interaction of Cyp33 and JARID1B *in vivo*

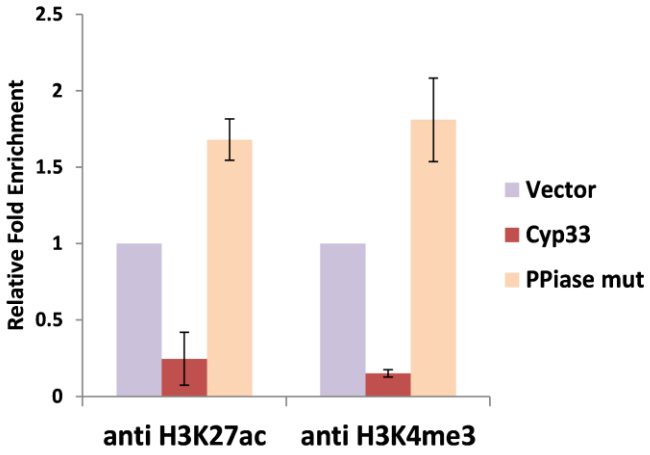
Co-Immunoprecipitation for FLAG JARID1B and HA Cyp33



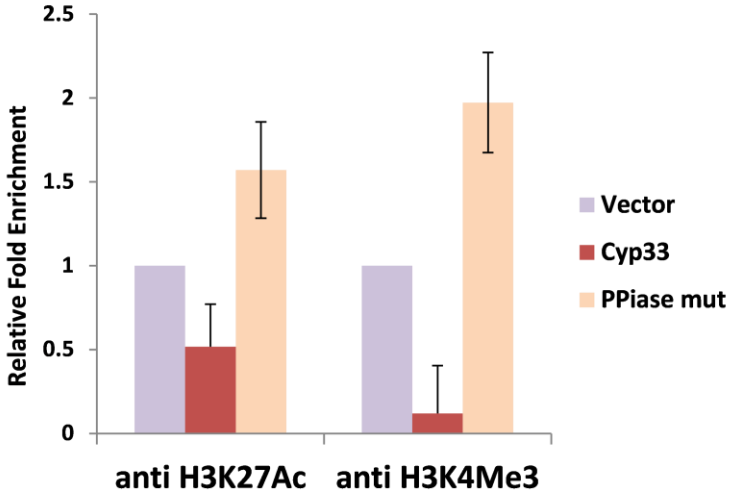
293T cells were transfected with indicated plasmids. After 48 hours cells were lysed in IPH lysis buffer. JARID1B was immuno-precipitated using an anti-FLAG antibody and immunoblotted for FLAG tagged JARID1B in the1st panel from the top, and for HA Cyp33 in the 2nd panel from the top. Third and fourth panels are inputs showing the over expression of FLAG tagged JARID1B and HA Cyp33 respectively. The fifth panel is an immunoblot for actin showing equal amount of protein is present in both samples. n=2, a representative experiment is shown.

Figure 11: Effect of PPIase activity of Cyp33 on H3K4Me3 and H3K27Ac levels and JARID1A and JARID1B recruitment at MLL target gene promoters.

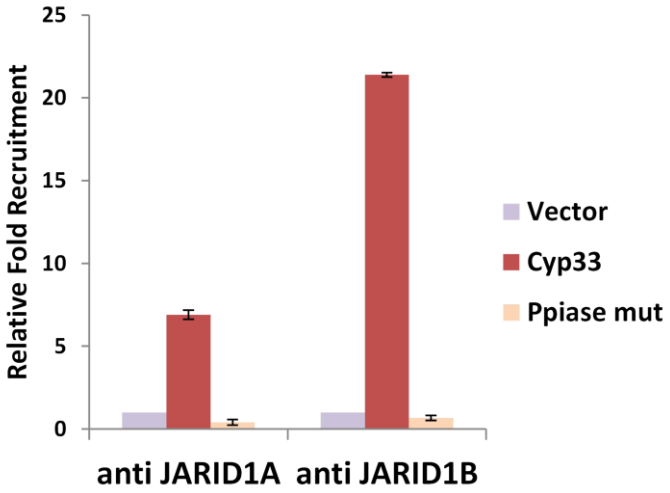
A. ChIP for H3K27Ac and H3K4Me3 at *HOXA9* promoter



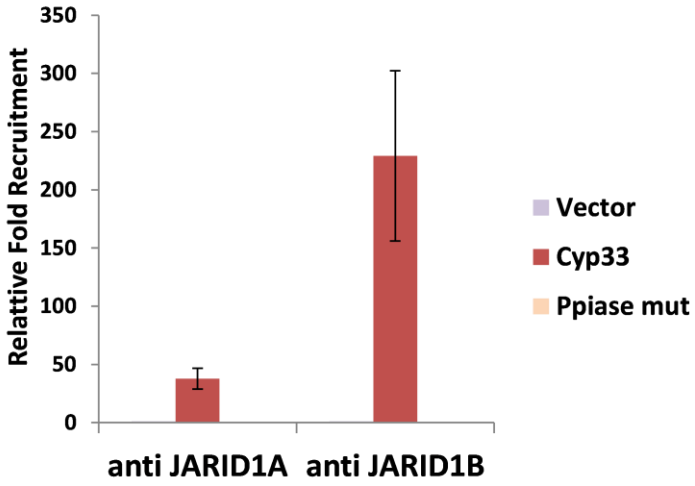
B. ChIP for H3K27Ac and H3K4Me3 at *CDKN1B* promoter



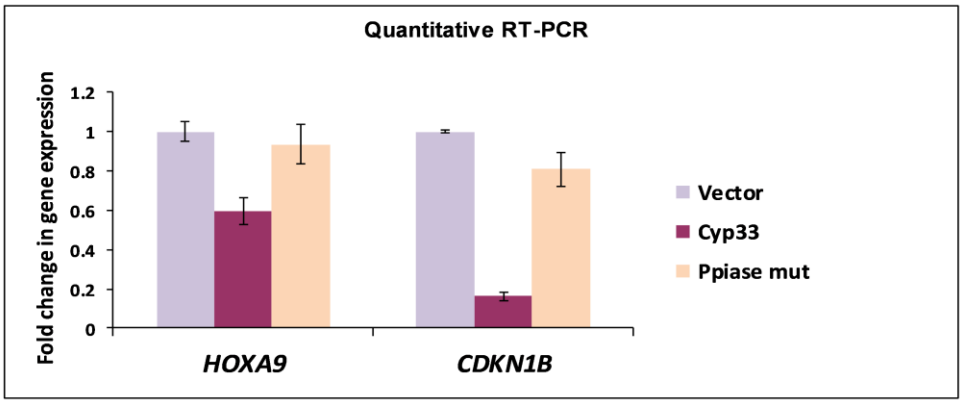
C. ChIP for JARID1A and JARID1B at *HOXA9* promoter



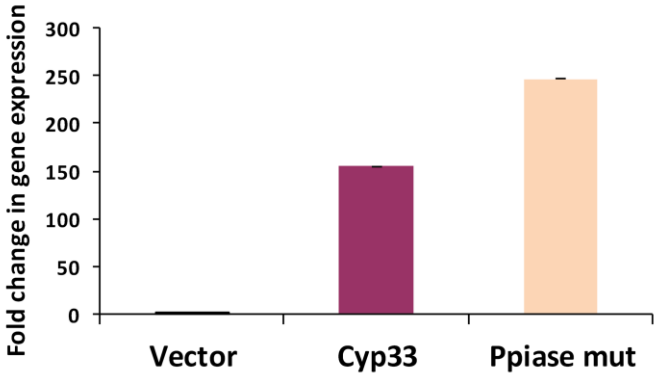
D. ChIP for JARID1A and JARID1B at *CDKN1B* promoter



E. qRT-PCR for *HOXA9* and *CDKN1B*



F. qRT-PCR for CYP33



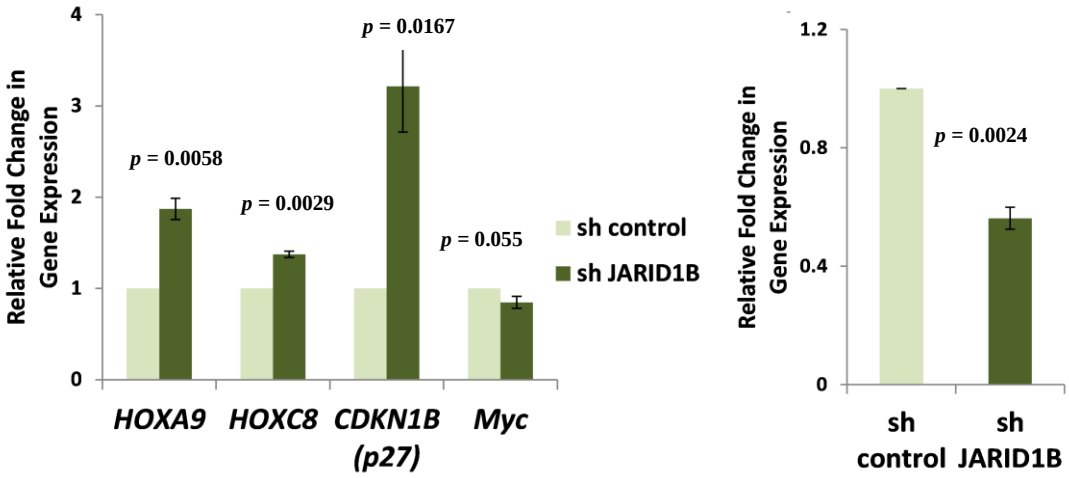
293T cells were transiently transfected with pCep4 (vector), or plasmids coding for pCep4-Cyp33, pCep4-Cyp33 PPIase mutant and harvested after 48 hours for RNA extraction and ChIP for H3K4Me3, H3K27Ac, JARID1A and JARID1B followed by q-PCR. **A,** ChIP for H3K4Me3 and H3K27Ac at *HOXA9* promoter; **B,** ChIP for H3K4Me3 and H3K27Ac at *CDKN1B* promoter; **C**. ChIP for JARID1A and JARID1B at *HOXA9* promoter; **D,** ChIP for JARID1A and JARID1B at *CDKN1B* promoter. Rabbit IgG was used as a negative control for the pull-down. n=1, error bar represents +/- S.D from the triplicates of q-PCR; **E**, qRT-PCR for *HOXA9* and *CDKN1B*; **F**, qRT-PCR for *CYP33*.n=1, error bar represents +/- S.D from the triplicates of q-PCR.

Experiments performed by Steven Poppen in our laboratory showed that the recruitment of transcriptional repressors like HDAC1 and Bmi-1 to the repression domain of MLL is PPIase dependent. JARID1A and JARID1B are present in complex with HDAC[103–105]. Therefore, we wanted to investigate if the PPIase activity of Cyp33 is required for Cyp33 mediated increased recruitment of JARID1A and JARID1B at MLL target gene promoters. Over expression of wild type Cyp33 resulted in dramatic increase in recruitment of JARID1A and JARID1B at MLL target promoters like *HOXA9* and *CDKN1B* when compared to vector control. Nevertheless, over expression of a mutant Cyp33 that results in 99% decrease in PPIase activity did not show the increased recruitment of JARID1A and JARID1B at MLL target promoters (Figure 11 C and D). Also, the PPIase mutant showed increased H3K4Me3 levels and H3K27Ac levels at the MLL target gene promoters (Figure 11A and B).

To determine if JARID1B is a transcriptional repressor of MLL target genes, shRNA mediated knock down of JARID1B was performed in 293T cells and the effect on the transcription of MLL target genes was measured by qRT-PCR (Figure 12A). The knock down of JARID1B was verified by qRT-PCR (Figure 12B)**.** About 50% decrease in the mRNA levels of JARID1B was achieved by shRNA. JARID1B knock down resulted in about two-fold increase in *HOXA9* and about 3.2 fold increase in *CDKN1B* transcription, whereas there was no effect on the transcription of other MLL target genes like *Myc* and *Meis1*. Thus JARID1B knock down affected the transcription of a sub-set of MLL target genes. The striking increase in JARID1B recruitment and a concomitant

Figure 12: Effect of knock down of JARID1B at MLL target genes

A. qRT-PCR for MLL target genes **B. qRT-PCR for *JARID1B***



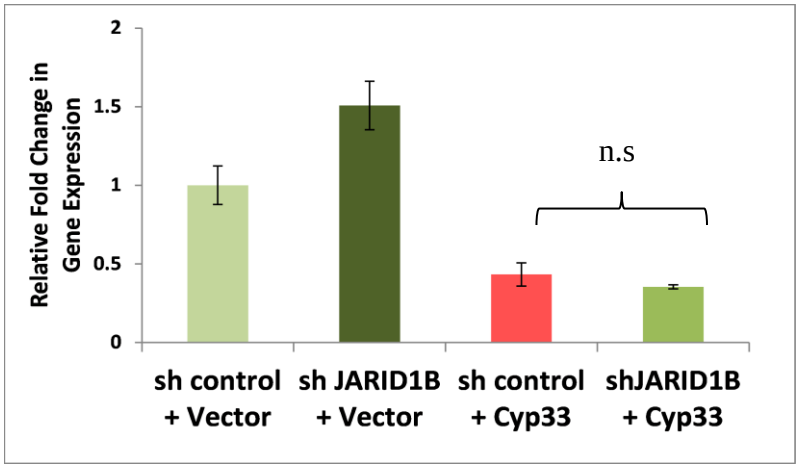
293T cells were stably transfected with plasmids for control shRNA or shRNA against JARID1B. Stable transfectants were selected in puromycin for 5 days. Cells were harvested for RNA extraction followed by q RT-PCR for **A**, MLL target genes; **B**, for *JARID1B*. n=3, average of 3 independent experiments and error bar represents S.D. *p* values are from two-tailed unpaired t-test.

decrease in H3K4Me3 levels at MLL target gene promoters upon Cyp33 over expression, suggested that JARID1B may be one of the downstream effectors of Cyp33.

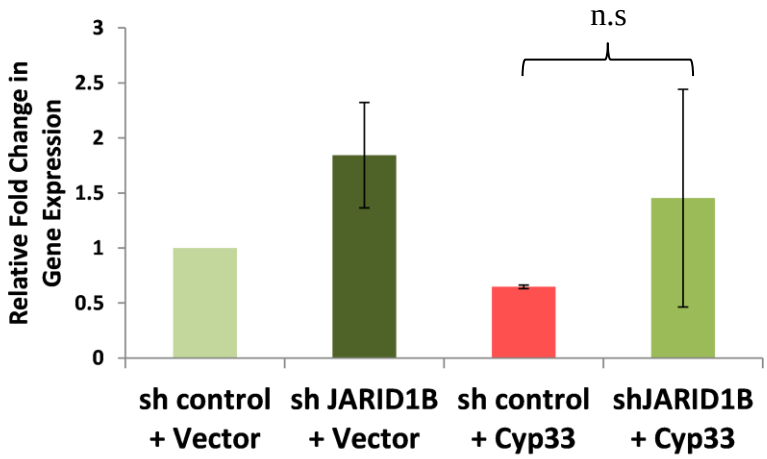
To know if JARID1B mediates Cyp33 transcription regulation of MLL target genes, Cyp33 was transiently over expressed in cells that have shRNA mediated stable knock down of JARID1B. pCep4 transfection after JARID1B knockdown resulted in statistically significant increase in expression of *HOXA9* (Figure 13A) and *CDKN1B* (Figure 13B) when compared to pCep4-Cyp33 transfection in sh control cells. This again shows that JARID1B is a transcriptional repressor of MLL target genes. Cyp33 over expression in sh control transfected cells resulted in a statistically significant decrease in transcription of *HOXA9* and *CDKN1B* when compared to sh control cells transfected with pCep4. This also reiterates that Cyp33 is a transcriptional repressor of MLL target genes. If JARID1B is a downstream effector of Cyp33 mediating transcriptional repression of MLL target genes, over expressing Cyp33 in the absence of JARID1B should not result in transcriptional repression of MLL target genes. However, Cyp33 over expression in JARID1B knock down cells resulted in statistically significant transcriptional repression of *HOX A9* and *CDKN1B* when compared to sh control cells that over express Cyp33 (Figure 13A and 13B). Hence, reducing the levels of JARID1B did not overcome Cyp33 mediated transcriptional repression of MLL target genes. Thus, even if JARID1B is a transcriptional repressor of MLL target genes; it may not be the only downstream effector of Cyp33. Over expression of Cyp33 and knock down of JARID1B was verified by qRT-PCR for *CYP33* and *JARID1B* (Figure 13C and 13D).

Figure 13: Effect of knock down of JARID1B and over expression of Cyp33 on the transcription of MLL target genes

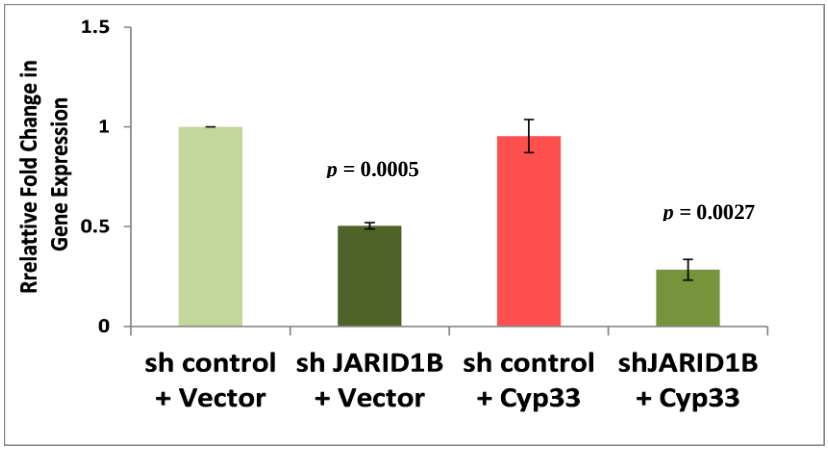
A. qRT-PCR for *HOXA9*



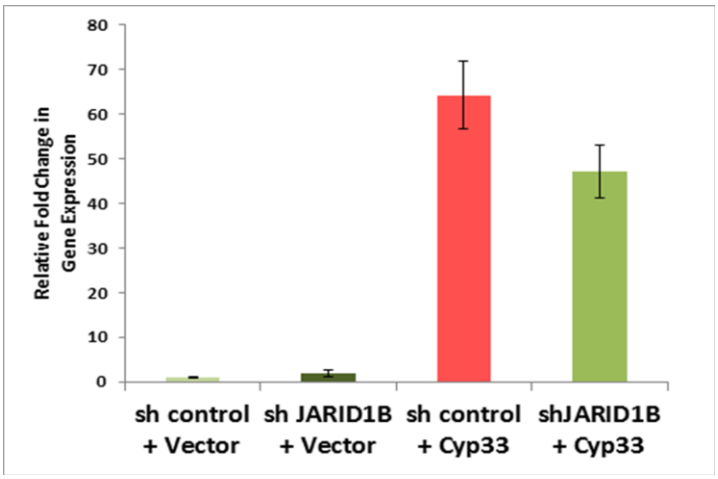
B. qRT-PCR for *CDKN1B*



C. qRT-PCR for *JARID1B*



D. qRT-PCR for *CYP33*



293T cells were stably transfected plasmids for control shRNA or shRNA against JARID1B. Stable transfectants were selected in puromycin for 5 days. Further cells were transfected with pCep4 (vector) or pCep4-Cyp33 plasmids. After 48 hours, cells were harvested for RNA extraction followed by qRT-PCR for: **A,** *HOXA9;* **B**, *CDKN1B*; **C**, *JARID1B;* and **D,** *Cyp33***.** n=2, average of 2 independent experiments and error bar represents S.D. n.s means not statistically significant.

Aim 2: Determine the significance of MLL PHD3 binding to H3K4Me3 vs Cyp33.

In addition to binding to Cyp33, the MLL PHD3 also binds to H3K4Me3. This binding was reported by us and other groups [26,29,30]. All these studies focused on the *in vitro* interactions of MLL PHD3 or Cyp33. Hence the biological significance of H3K4Me3 binding and Cyp33 binding to the MLL PHD3 is yet to be determined. From all these studies we know which are the amino acids in the MLL PHD3 finger that contribute to H3K4Me3 binding and Cyp33 binding. Conserved aromatic and hydrophobic residues in the PHD3 of MLL form a cage that acts as a binding pocket for H3K4Me3 [26,29,30]. Chang et al. have shown that the M1585A mutation abrogates H3K4Me3 binding and resulted in decreased transcriptional activation of MLL target genes. Unlike the W1594A point mutant, this mutation did not affect the protein stability [30]. Based on this information, we mutated methionine 1585 to alanine (M1585A) as a candidate mutation to disrupt the binding to H3K4Me3. Wang et al. have shown that mutating methionine 1606 to aspartate in MLL abolished the binding of MLL PHD3 to Cyp33 [29]. Moreover, methionine 1606 does not participate in H3K4Me3 binding. Hence M1606D mutation should not interfere with H3K4Me3 binding and was used as a candidate mutation that disrupts MLL PHD3 binding to Cyp33.

MLL is required for normal hematopoiesis [6,7,67,106]. In order to understand the biological significance of MLL PHD3 binding to H3K4Me3 versus Cyp33, we wanted to study the effect of the mutations in the context of a full length MLL protein in hematopoietic progenitor colony formation assay. But MLL is a big protein and the cDNA is about 14KB long. Therefore, it will be difficult to mutate, clone and will be impossible to package such a big construct in retro viruses. MLL is proteolytically

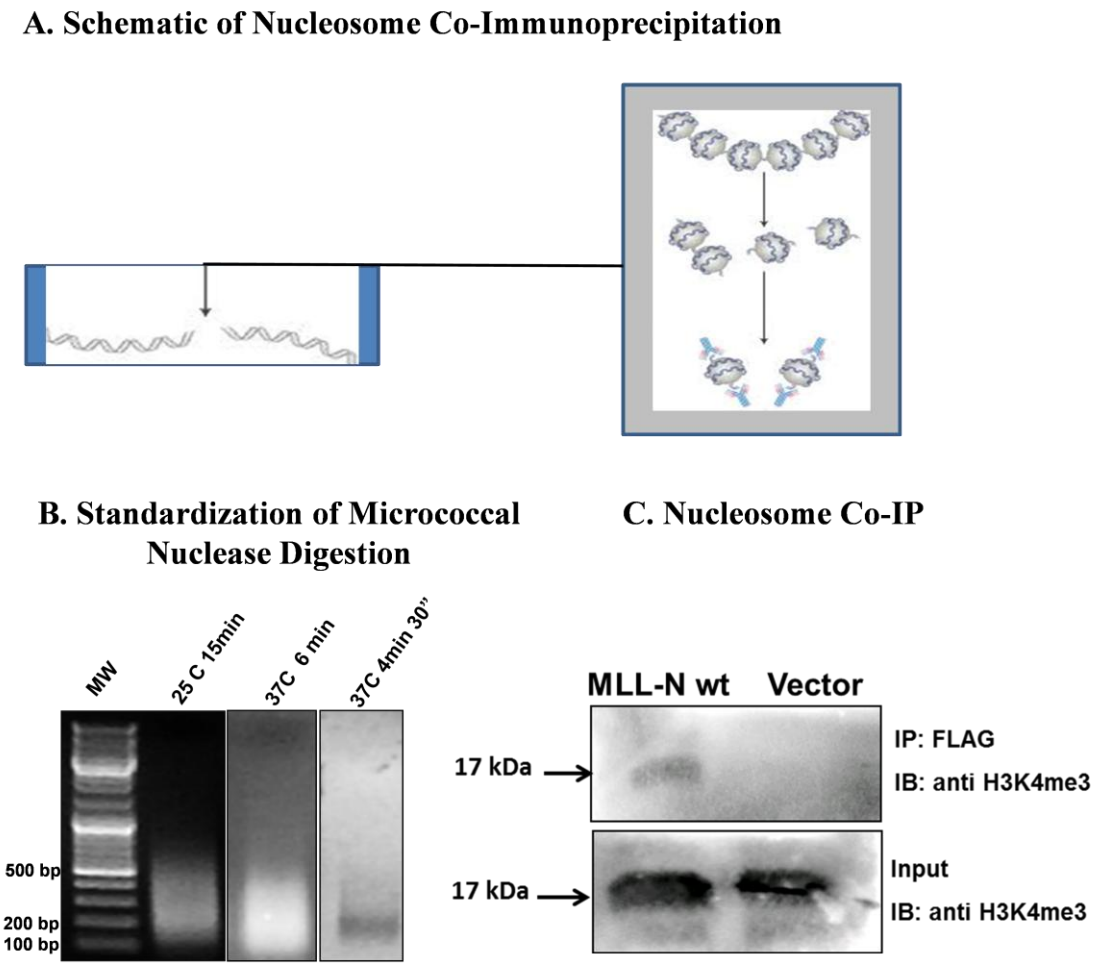
cleaved into MLL-N and MLL-C; the fragments interact with each other through their FYRN and FYRC domains, and MLL-C stabilizes MLL-N [36,37,74]. The N-fragment has the PHD3 finger that binds to H3K4Me3 and the C-fragment has the H3K4Me3 methyl transferase activity. To circumvent the practical problems of expressing a big construct, the mutations were made in a DNA sequence encoding the MLL-N fragment instead of full length MLL. FLAG tagged MLL-N from the pCI-MLL-N provided by Dr. James Hsieh was cloned into MSCV neo by Wei Wei in our laboratory. MSCV neo FLAG MLL-N M185A and M1606D were generated in this study by site-directed mutagenesis in a smaller fragment of MLL-N, and swapped into MSCV MLL-N wt as explained in the Methods section.

To date, the interaction of the MLL PHD3 with H3K4Me3 has been only demonstrated *in vitro*. To determine if MLL-N wt (which has the PHD3) binds to H3K4Me3 *in vivo*, a nucleosome pull down assay was performed. If MLL-N (through its PHD3) binds to H3K4Me3, then it would co-immunoprecipitate with nucleosomes that are trimethylated at H3K4. Partial digestion of the chromatin fraction with micrococcal nuclease will result in induction of double stranded breaks of DNA in the nucleosome linker region, releasing mononucleosomes**.** Enrichment for the mononuclesome fraction was optimized by varying the time and temperature of digestion. To standardize the micrococcal nuclease assay, chromatin was isolated from 293T cells as explained in methods section and the micrococcal nuclease digestion was performed at the indicated time and temperature (Figure 14B)**.** About 146 bp of DNA are wound around a mononucleosome. To verify mononucleosome enrichment, the DNA was purified from proteins and loaded onto a 2% agarose gel containing ethidium bromide. DNA fractions,

with an average length of about 150 bp were obtained by performing the digestion at 37°C for 4 minutes and 30 seconds (Figure 14B).

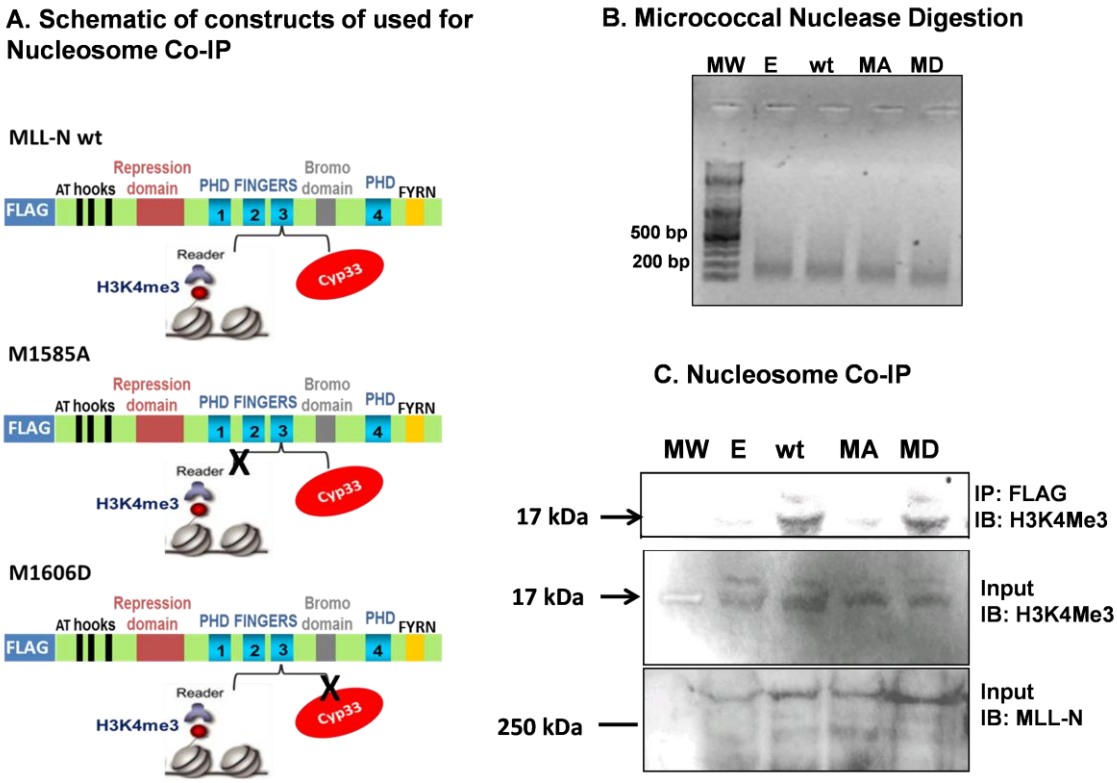
A nucleosome Co-IP of FLAG MLL-N and H3K4Me3 was performed (Figure 14C). FLAG MLL-N co-immunoprecipitated with trimethylated H3K4, whereas the vector control did not. This shows that there is no non-specific binding of the nucleosomes to the protein A conjugated FLAG beads. To ensure that there is equal amount of H3K4Me3 containing nucleosomes present in the nucleosome preparations, a Western blot for H3K4Me3 was performed (Figure 14C). After establishing the conditions of the nucleosome Co-IP with MLL-N wt, the binding status of M1585A and M1606D mutants to H3K4Me3 was determined using the same assay. The micrococcal digestion and the release of mononucleosomes were verified (Figure 15B)**.** A nucleosome Co-IP of FLAG MLL-N or the mutants with H3K4Me3 performed (Figure 15C) showed that both MLL-N wt and the M1606D mutant co-immunoprecipitated with nucleosomes containing H3K4Me3. In contrast, the M1585A mutant behaved similar to that of the vector control (negative control). Hence a M1585A mutation in the PHD3 of MLL-N abrogates binding to H3K4Me3 marked nucleosomes.

Figure 14: Verification of binding of MLL-N wt to H3K4Me3 nucleosomes *in vivo*



A, Schematic of Nucleosome pull-down assay; **B**, Standardization of micrococcal nuclease digestion was performed in untransfected 293T cells. Enrichment for mononucleosomes was verified by agarose gel, n=1; **C,** Nucleosome Co-ip, 293T cells were transiently transfected with FLAG MLL-N wt, cells or MSCV neo (Vector) plasmids. Cells were harvested after 48 hrs for micrococcal digestions at 37C for 4 min 30" followed by nucleosome pull down assay. FLAG tagged MLL-N was immunoprecipitated with anti-FLAG agarose beads and immunoblotted for H3K4 trimethylated nucleosomes with anti H3K4Me3 antibody, n=2, a representative experiment is shown.

Figure 15: Binding status of M1585A and M1606D to H3K4Me3 nucleosomes *in vivo*



293T cells were transfected with MSCV neo (E) or FLAG tagged MLL-N wt (wt), M1585A (MA), M1606D (MD) plasmids. Cells were harvested after 48 hrs for a nucleosome pull down assay: **A,** Schematic of binding of MLL-N wt and mutants to H3K4Me3 and Cyp33; **B**, Micrococcal nuclease digestion performed to enrich for mononucleosomes which was verified by DNA agarose gel, (MW means DNA molecular weight marker); **C,** Nucleosome Co-IP Nucleosome Co-IP: FLAG tagged MLL-N was immunoprecipitated with anti-FLAG agarose beads and immunoblotted for H3K4 trimethylated nucleosomes with anti H3K4Me3 antibody. MW means n=2, a representative experiment is shown.

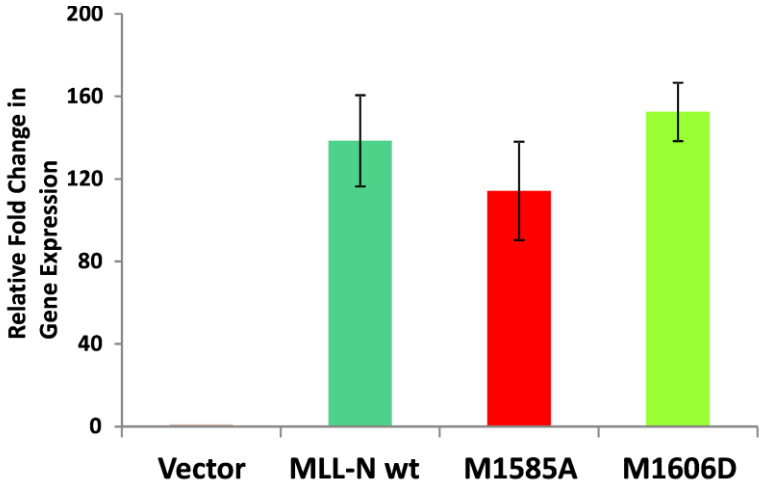
The binding status of MLL-N wt, or the mutants, to Cyp33 was determined using a GST pull down experiment, by Dr. Osmers in our laboratory. As expected the MLL-N wt and M1585A mutant both bind Cyp33, but the M1606D mutation does not.

H3K4Me3 is a transcriptional activation mark. Hence, MLL PHD3 binding to H3K4Me3 may protect H3K4Me3 from being targeted by histone demethylases. Furthermore, MLL is a writer of H3K4Me3 marks. Hence, MLL PHD3 binding to H3K4Me3 may help MLL-C in the propagation of methyl marks to adjacent nucleosomes. Both these scenarios should result in an increase in H3K4Me3 marks at MLL target gene promoters. As a consequence, MLL PHD3 binding to H3K4Me3 should result in transcriptional activation of MLL target genes and abrogating the binding should result in decrease in transcription of MLL target genes.

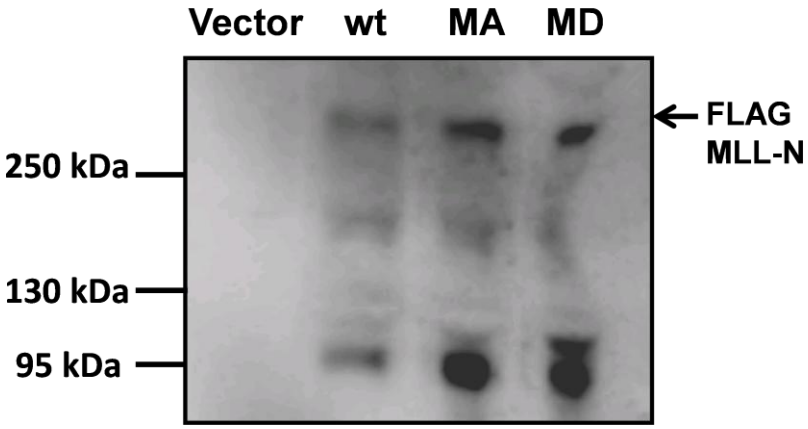
From *in vitro* binding studies, it is known that Cyp33 has a higher affinity for MLL PHD3 than H3K4Me3 and Cyp33 in excess reduces the affinity of MLL PHD3 for H3K4Me3 [26]. Cyp33 over expression in our studies also resulted in a decrease of H3K4Me3 marks and increased recruitment of JARID1B histone demethylases at MLL target gene promoters. Hence abrogating Cyp33 binding in MLL PHD3 should reduce the exposure of H3K4Me3 to the JARID1B histone demethylase, resulting in an increase in H3K4Me3 marks and increase in transcription of MLL target genes. In short, the hypothesis of the second aim of this study is that MLL PHD3 binding to H3K4Me3 contributes to transcriptional activation, while MLL PHD3 binding to Cyp33 contributes to transcriptional repression of MLL target genes.

Figure 16: Verification of over expression of MLL-N

A. qRT-PCR for *MLL*



B. Western Blot for FLAG tagged MLL-N and Mutants



MEFs (MLL +/+) were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D (in the presence of endogenous MLLC) and selected with G-418 for 11 days. Cells were harvested for RNA followed by: **A**, qRT-PCR for human *MLL*. n=3, average of 3 independent experiments; **B**, Immunoblot for FLAG tagged MLL-N and mutants. n=2, a representative experiment is shown.

To test the above mentioned hypothesis, MEF (*Mll+/+*) were transduced with MSCV neo (Vector), MLL-N wt., M1585A or M1606D expressing retro-viruses and the effect on target gene expression was assessed by qRT-PCR. Over expression of *MLL-N* wt and mutants was verified by qRT-PCR (Figure 16A). *MLL* transcripts were normalized to *Actb* transcripts levels. There is about the same level of over expression of MLL-N wt and mutants. Over expression of MLL-N at the protein level was verified by immunoblot for FLAG tag (Figure 16B).

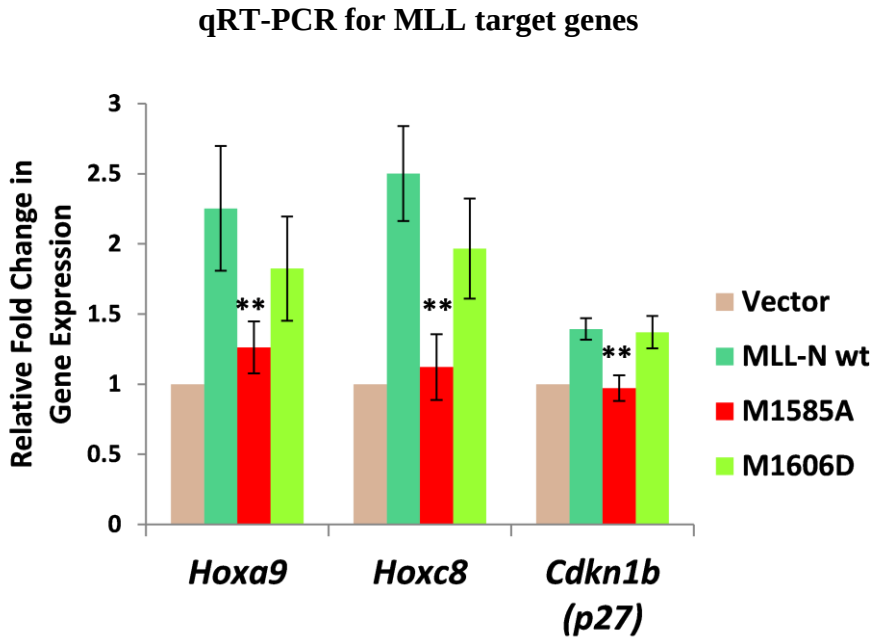
The effect of over expression of MLL-N wt and mutant on target gene expression was measured by qRT-PCR (Figure 17). Over expression of MLL-N wt resulted in a statistically significant increase in transcription of the MLL target genes like *Hoxa9, Hoxc8,* and *Cdkn1b* when compared to the vector control. The *p* values from one sample Student's t test when compared between vector and MLL-N wt for *Hoxa9, Hoxc8* and *Cdkn1b* are 0.0068, 0.0097, and 0.0087 respectively. Over expression of M1606D mutant (that does not bind Cyp33) resulted in transcriptional activation of *Hoxa9, Hoxc8* and *Cdkn1b,* a trend similar to that of MLL-N wt over expression. However, the *p* values from one sample student's t test when compared between vector and M1606D for *Hoxa9, Hoxc8* and *Cdkn1b* are 0.0646, 0.0204 and 0.0495 showed that the statistically significant transcriptional activation was observed only for *Hoxc8* and *Cdkn1b* but not for *Hoxa9.* Additionally, the M1606D mutation did not result in increased transcription of MLL target genes over the MLL-N wt over expression.

In contrast to MLL-N wt and M1606D mutant, over expression of MLL-N M1585A (that does not bind H3K4Me3) did not have any effect on transcription of MLL target genes when compared to the vector control. However, M1585A over expression

resulted in a decrease in transcription of MLL target genes when compared to MLL-N wt and its M1606D mutant. The *p* values from two-tailed unpaired student's t-test compared between M1585A and MLL-N wt for *Hoxa9*, *Hoxc8* and *Cdkn1b* are 0.0063, 0.0025 and 0.0037 respectively. The *p* values from two-tailed unpaired student's t-test compared between M1585A and M1606D for *Hoxa9*, *Hoxc8* and *Cdkn1b* are 0.0352, 0.012 and 0.0094 respectively. This shows that MLL PHD3 binding to H3K4Me3 contributes to the transcriptional activation of MLL target genes.

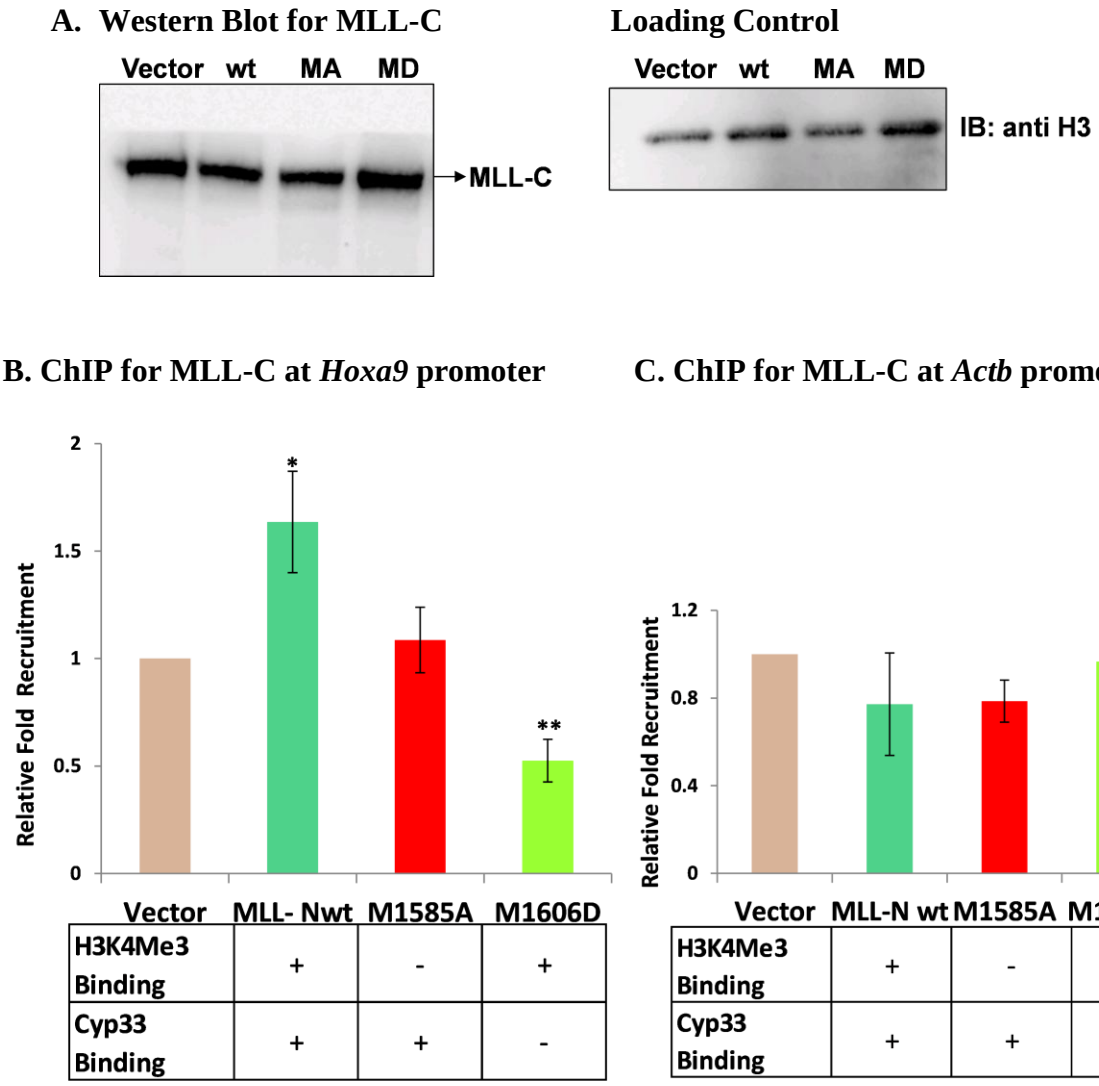
Since MLL-N wt over expression resulted in transcriptional activation of MLL target genes, I hypothesized that MLL-N over expression resulted in increased protein stability of MLL-C or nuclear import of MLL-C. To test this, MEFs were transduced with MSCV neo (vector), or MLL-N wt, M1585A and M1606D retro-viral expression constructs and cells were harvested for nuclear lysate preparation. Further, a Western blot for MLL-C (Figure 18A) showed that MLL-N wt over expression did not increase MLL-C protein levels. Alternatively, I hypothesized that MLL-N wt over expression increased the recruitment of MLL-C at MLL target gene promoters, thus leading to increase in transcription of MLL target genes. Hence a ChIP for MLL-C at the *Hoxa9* promoter was performed in (Figure 18B). MLL-N wt over expression resulted in a statistically significant increase in MLL-C recruitment at the *Hoxa9* promoter when compared to the vector control (one sample student's t-test $p = 0.031$), M1585A (unpaired student's t- test $p = 0.0017$ and M1606D (unpaired student's t- test $p = 0.0322$). However there is no increase in MLL-C recruitment upon MLL-N wt over expression at the house-keeping gene promoter, *Actb* promoter (Figure 18C). Unlike MLL-N wt over expression, M1585A over expression did not result in increase in MLL-C recruitment at the *Hoxa9*

Figure 17: Effect of MLL-N wt and mutants over expression on transcription of MLL target genes



MEFs (MLL +/+) were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D retroviruses (in the presence of endogenous MLLC) and selected with G-418 for 11 days. Cells were harvested for RNA preparation followed by quantitative RT-PCR for MLL target genes. n=3, average of 3 independent experiments, error bars represent S.D. Pairwise comparisons between M1585A and MLL-Nwt ; M1585A and M1606D were performed by two-tailed unpaired Student's t-test. ** - Decrease in transcription of MLL target genes in M1585A is statistically significant when compared to MLL-N wt (*Hoxa9- p* = 0.0063, *Hoxc8- p* = 0.0025, *Cdkn1b- p* = 0.037) and M1606D (*Hoxa9- p* = 0.0352, *Hoxc8- p* = 0.012, *Cdkn1b- p* = 0.0094).

Figure 18: Effect of MLL-N and mutant over expression on MLL-C recruitment



MEFs (MLL +/+) were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D retroviruses (in the presence of endogenous MLLC) and selected with G-418 for 11 days. Cells were harvested for: **A,** Immunoblot for MLL-C, Cells were lysed for nuclear lysate preparation, followed by immunoblot for MLL-C, n=1; **B**, ChIP for MLL-C at the *Hoxa9* promoter. Pairwise comparisons were performed by either one sample or unpaired two-tailed student's t-tests. * means increase in MLL-C recruitment in MLL-N wt when compared to vector (one sample student's t test $p = 0.031$) and M1585A (unpaired student's t- test $p = 0.0017$). ** means decrease in MLL-C recruitment in M1606D when compared to MLL-N wt (unpaired student's t- test $p = 0.0322$) and to M1585A (unpaired student's t- test $p = 0.0011$). **C**, ChIP for MLL-C at *Actb* promoter. n=3, average of 3 independent experiments, error bar represents S.D.

promoter when compared to vector control. The *p* value from one sample student's t-test performed between vector control and MLL-N wt is 0.1421. This shows that MLL PHD3 binding to H3K4Me3 may be involved in the recruitment of MLL-C to *Hoxa9* promoter. Surprisingly, the M1606D mutant (that does not bind Cyp33) showed a decrease in recruitment of MLL-C at the *Hoxa9* promoter when compared to, M1585A (unpaired student's t- test $p = 0.0011$) and MLL-N wt (unpaired student's t- test $p = 0.0322$).

Since over expression of MLL-N wt and M1606D mutant resulted in transcriptional activation of MLL target genes but M1585A did not, we hypothesized that MLL PHD3 binding to H3K4Me3 contributes to transcriptional activation of MLL target genes. This may be because, MLL PHD3 binding to H3K4Me3 protects H3K4Me3 from being demethylated by histone demethylases and/or helps in propagation of methyl marks to adjacent nucleosomes. Both the above mentioned scenarios should result in increased H3K4Me3 at MLL target genes resulting in increased transcriptional activation of MLL target genes. To calculate the fold enrichment for H3K4Me3 at a given promoter, the amount of H3 present at a given promoter should be determined. A ChIP for histone H3 at MLL target genes was performed (Figure 20A). MLL-N wt and mutant over expression did not significantly affect the total H3 levels both at MLL target genes and *Actb* promoter (Figure 19 A & B).

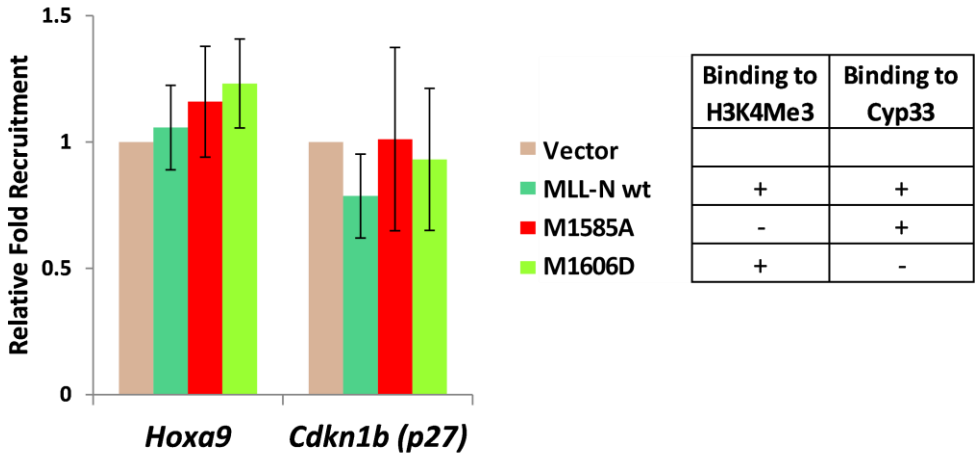
After determining the histone H3 density at MLL target genes, a ChIP for H3K4Me3 at MLL target gene promoters was performed (Figure 20A). Over expression of MLL-N wt resulted in an increase in H3K4Me3 marks relative to vector control at the *Hoxa9* promoter, but did not show a statistically significant increase at the *Cdkn1b* promoter. The *p* values from one sample student's t-test are 0.0067 and 0.1187

respectively. Over expression of MLL-N wt also resulted in a statistically significant increase relative to the M1585A mutant in H3K4Me3 marks at the *Hoxa9* (unpaired student's t-test p =0.001) and *Cdkn1b* (unpaired student's t-test p =0141) promoters but not at the *Actb* promoter (Figure 20B). Abrogating H3K4Me3 binding resulted in a decrease in H3K4Me3 marks, a decrease in MLL-C recruitment and decrease in target gene expression. These results together suggest that MLL PHD3 binding to H3K4Me3, helps MLL-C to methylate adjacent nucleosomes. The M1606D mutant did not show a significant increase in H3K4Me3 marks over the vector control and the M1585A mutant. This may be a result of decreased recruitment of MLL-C at the MLL target promoters (Figure 20A).

Next, to determine if MLL PHD3 binding to H3K4Me3 protects H3K4Me3 from histone demethylases, MLL-N wt or its mutants were over expressed, and recruitment of the JARID1B histone demethylase at the MLL target gene promoters was measured by ChIP. If MLL PHD3 binding to H3K4Me3 protects H3K4Me3 from histone demethylases, then abrogating MLL PHD3 binding to H3K4Me3 should result in increased recruitment of JARID1 histone demethylases at MLL target gene promoters. There was no significant difference between vector control, MLL-N wt and M1585A in the recruitment of JARID1B at *Hoxa9* promoter (Figure 21A). This shows that in the system used, MLL PHD3 binding to H3K4Me3 does not protect H3K4Me3 from JARID1B binding. M1606D over expression resulted in a statistically significant decrease in the recruitment of JARID1B at the *Hoxa9* promoter when compared to the vector control (one sample t-test p = 0.0046, MLL-N wt (unpaired student's t-test p = 0.034) or M1585A (unpaired student's t-test p = 0.0351).

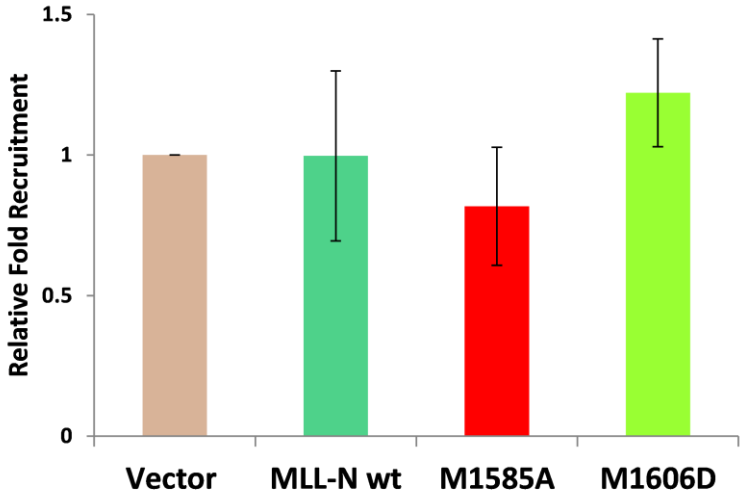
Figure 19: Effect of over expression of MLL-N wt and mutants on histone H3 density

A. ChIP for Histone H3 at MLL target gene promoters.



Binding to H3K4Me3	Binding to Cyp33
+	+
-	+
+	-

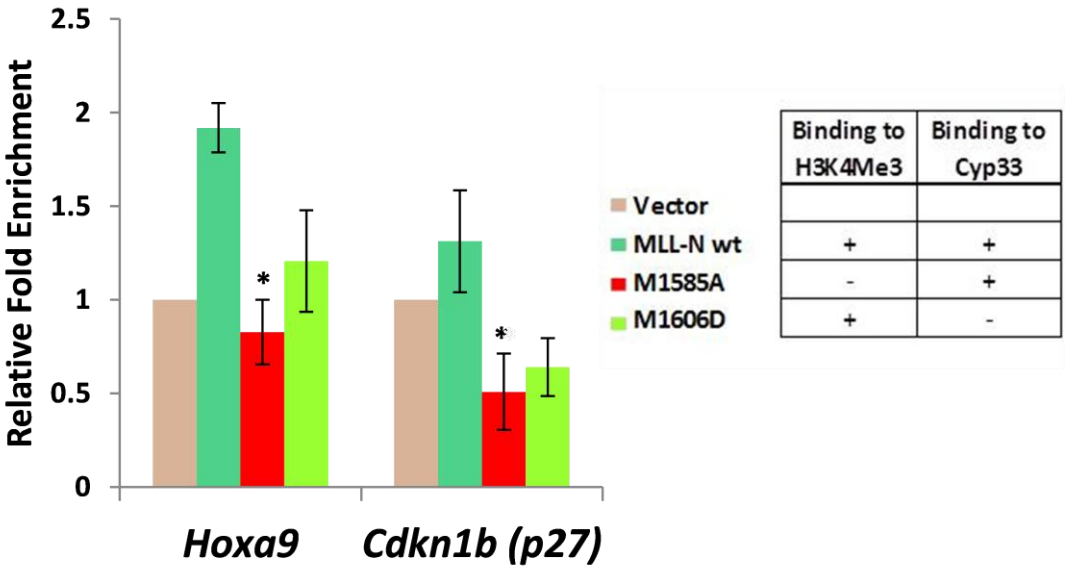
B. ChIP for Histone H3 at *Actb* promoter.



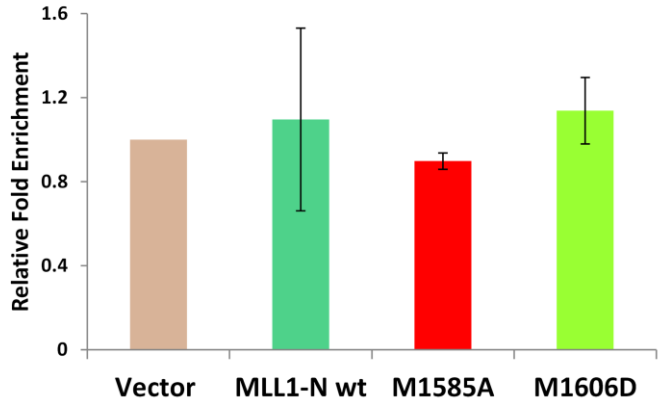
MEFs (MLL +/+) were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D (in the presence of endogenous MLLC) and selected with G-418 for 11 days and harvested for ChIP for histone H3 at: **A**, MLL target gene promoters, **B**, *Actb* promoter was performed. n=3, average of 3 independent experiments, error bars represent S.D.

Figure 20: Effect of over expression of MLL-N wt and mutants on H3K4Me3 marks

A. ChIP for H3K4Me3 at MLL target gene promoters



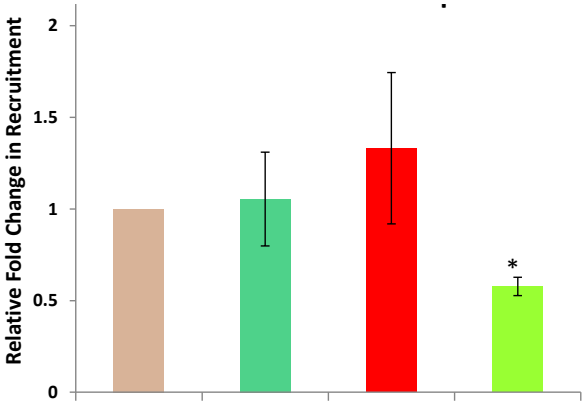
B. ChIP for H3K4Me3 at *Actb* promoter.



MEFs (MLL +/+) were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D (in the presence of endogenous MLLC) and selected with G-418 for 11 days and harvested for ChIP for histone H3K4Me3 at **A**, MLL target gene promoters, **B**, *Actb* promoter in figure B was performed. n=3, average of 3 independent experiments, error bars represent S.D. * Decrease in H3K4Me3 levels in M1585A at the *Hoxa9* (unpaired student's t-test p =0.001) and *Cdkn1b* (unpaired student's t-test p =0141) when compared to MLL-N wt.

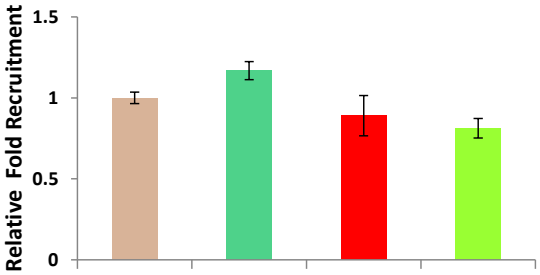
Figure 21: Effect of over expression of MLL-N wt and mutants on JARID1B recruitment

A. ChIP for JARID1B at *Hoxa9* promoter



	Vector	MLLN-wt	M1585A	M1606D
H3K4Me3 Binding	+	+	-	+
Cyp33 Binding	+	+	+	-

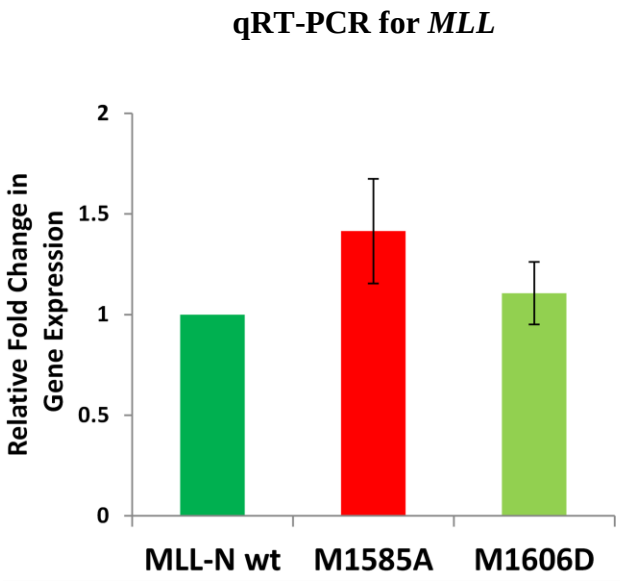
B. ChIP for JARID1B at *Actb* promoter



	Vector	MLL-N wt	M1585A	M1606D
H3K4Me3 Binding	+	+	-	+
Cyp33 Binding	+	+	+	-

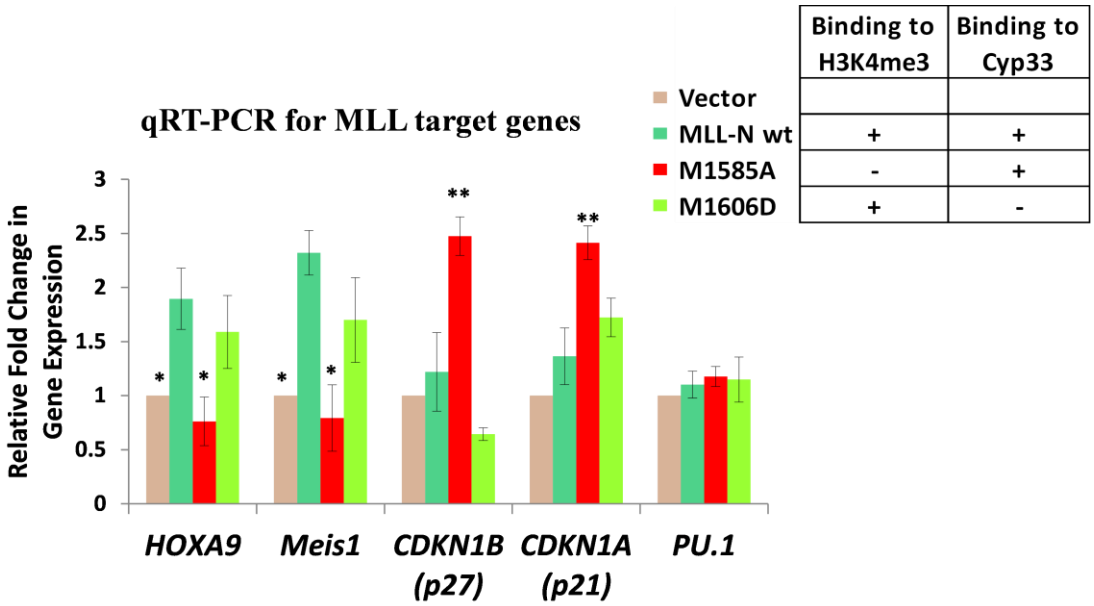
MEFs (MLL +/+) were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D (in the presence of endogenous MLLC) and selected with G-418 for 11 days and harvested for ChIP for: **A**, JARID1B at MLL target gene promoters. *- Decrease in JARID1B recruitment when compared to vector control (one sample t-test $p = 0.0046$) MLL-N wt- $p = 0.034$ and M1585A- $p = 0.0351$; **B**, JARID1B at *Actb* promoter. n=3, average of 3 independent experiments, error bars represent S.D.

Figure 22: Verification of over expression of MLL-N wt and mutants in mouse hematopoietic progenitors



C kit + mouse hematopoietic progenitors isolated from mouse bone marrow were retro virally transduced with MSCV neo (Vector), MLL-N wt, M1585A, and M1606D and grown in RPMI containing IL-3, IL-6, SCF, GM-CSF and selected with G418 for 3 days. The cells were harvested for RNA extraction followed by qRT-PCR for human *MLL*. n=3, average of 3 independent experiments, error bars represent S.D.

Figure 23: Effect of over expression of MLL-N wt and mutants on MLL target gene expression in Ckit+ve mouse hematopoietic progenitors



Transduction and selection as explained in Figure. Cells were harvested for RNA extraction. qRT-PCR for MLL target genes was performed. *- Statistically significant decrease in expression of *Hoxa9* and *Meis1* in Vector and M1585A when compared to MLL-N wt over expression. ** Statistically significant increase in expression of *Cdkn1b* and *Cdkn1a* in M1585A when compared to Vector, M1585A and M1606D. n=3, average of 3 independent experiments, error bars represent S.D.

Hence, abrogating MLL PHD3 binding to Cyp33 (M1606D) resulted in decreased recruitment of histone demethylases at MLL target gene promoters whereas at the *Actb* promoter there was no significant decrease.

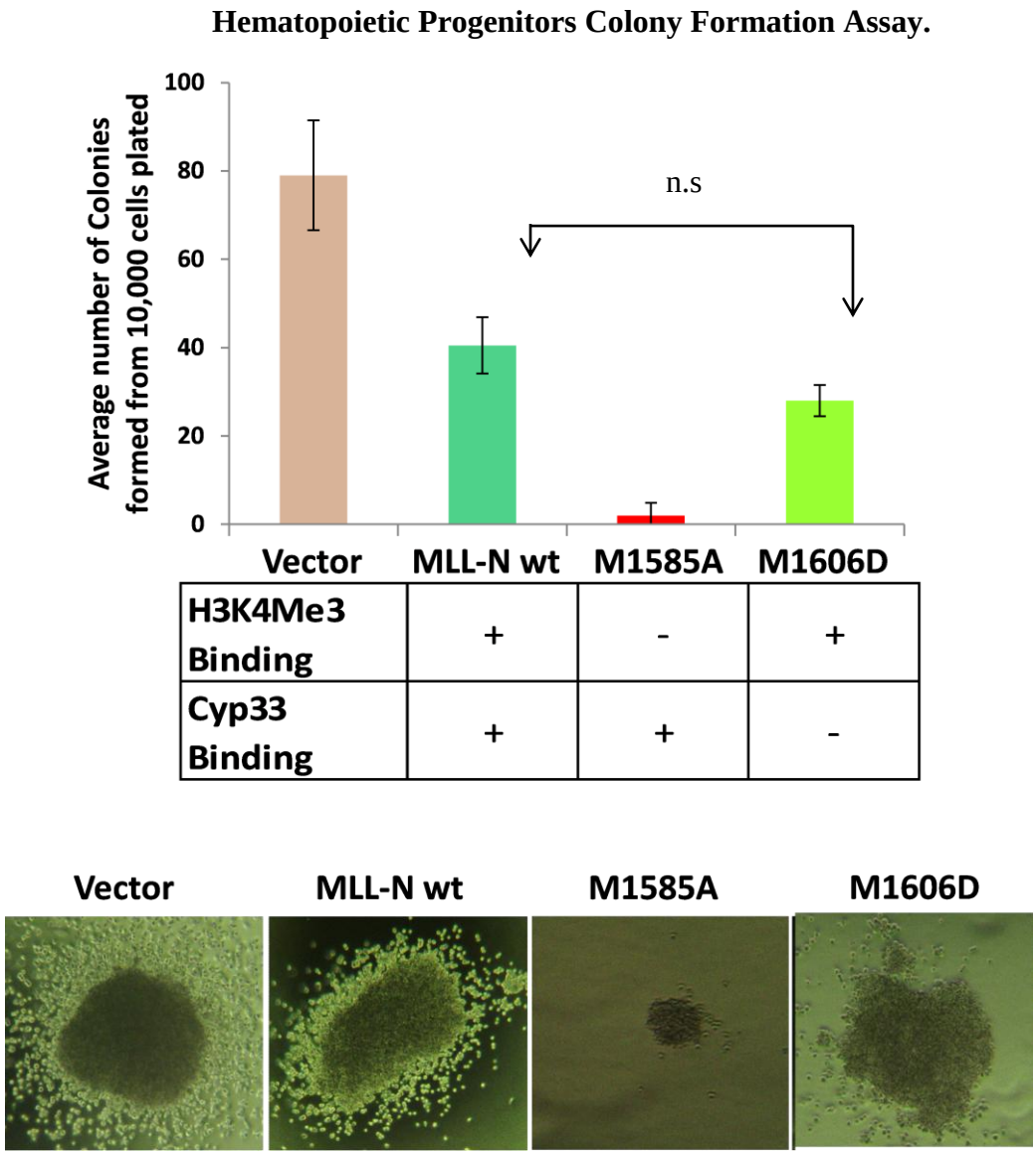
Since over expression of MLL-N wt or mutants had an effect on MLL target genes in MEFs, I wanted to determine the effect of over expression of MLL-N wt and its mutants in mouse hematopoietic progenitors. Over expression of MLL-N wt, its mutants M1585A and M1606D in C-kit+ mouse hematopoietic progenitors after retro viral transduction was verified by qRT-PCR for human *MLL* normalized to *Actb*. The data in Figure 23 shows that MLL-N wt, M1585A and M1606D are expressed at about equal levels in hematopoietic progenitors.

Having confirmed that the retro-viral transductions have worked in hematopoietic progenitors the effect of over expression of MLL-N wt and mutants on target gene expression was assessed by qRT-PCR (Figure 23). Similar to MEF, over expression of MLL-N wt and M1606D in hematopoietic progenitors resulted in statistically significant increase in expression of *Hoxa9* when compared to vector and M1585A. The *p* value, from Student's t-test, when comparing between MLL-N wt and M1585A over expression is 0.0056. The *p* value for M1606D and M1585A is 0.0243. Over expression of MLL-N wt and M1606D also resulted in increased transcription of Meis1 when compared to vector and M1585A. This shows that MLL PHD3 binding to H3K4Me3 in hematopoietic progenitors contributes to transcriptional activation of MLL target genes.

Like *Hoxa9* and *Meis1*, *Cdkn1a* and *Cdkn1b* are known targets of MLL [10,11]. In contrast to *Hoxa9* and *Meis1*, over expression of MLL-N wt did not increase the

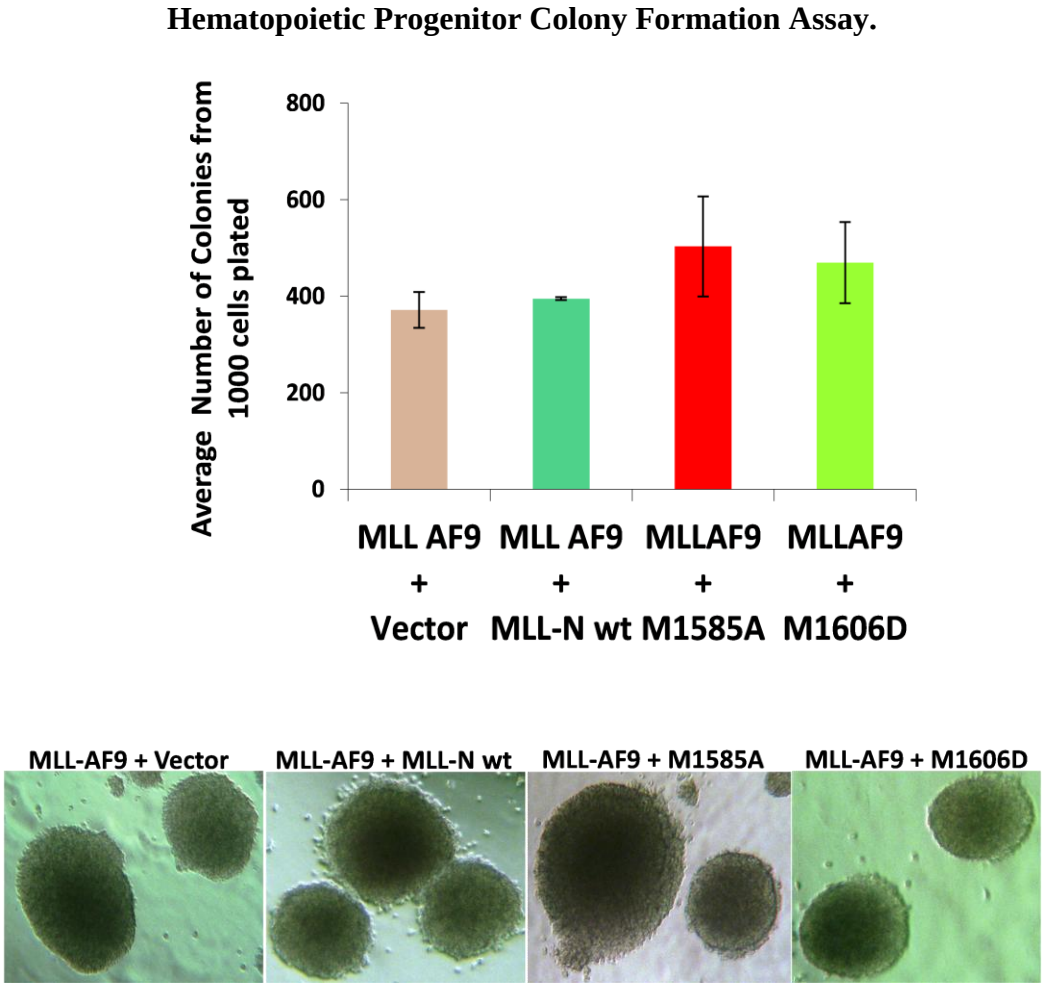
transcription of *Cdkn1b* and *Cdkn1a*. Whereas the over expression of M1585A resulted in increased expression of *Cdk1b* and *Cdkn1a*. The expression of *Pu.1* remained unaffected by MLL-N and mutant over expression. *Hoxa9* and *Meis1* expression keeps the hematopoietic progenitors in an undifferentiated state. Upon differentiation their expression decreases. Since MLL PHD3 binding to H3K4Me3 contributes to transcriptional activation of *Hoxa9* and *Meis1*, we tested the effects of MLL-N wt on hematopoiesis using an *in vitro* colony formation assay. If the cells are in undifferentiated state, each progenitor would undergo multiple cell divisions giving rise to a colony in methyl cellulose. This assay is a measure of the proliferation potential of hematopoietic progenitors. A portion of transduced Ckit+ mouse hematopoietic progenitors used in the previous experiment (Figure 23) were plated on methyl cellulose containing IL-3, IL-6, GM-CSF and SCF. The successful transduction and over expression of MLL-wt and mutants was verified (Figure 22). When compared to vector control transduced cells, MLL-N over expression resulted in a decrease in colony formation. Considering the fact that the MSCV neo MLL-N retroviral expression construct is at least twice longer than the MSCV neo vector, the differences in the viral titer and expression of neo-resistance will affect the colony formation. Since MLL-N wt, M1585A and M1606D are the same molecular weight and they have about equal levels of expression, I compared the colony formation potential of MLL-N wt to that of the mutants. M185A over expression resulted in a drastic reduction in colony formation when compared to MLL-N wt and vector control (Figure 24); over expression of the M1606D mutant showed no difference with MLL-N wt. Hence MLL PHD3 binding to H3K4Me3 contributes to MLL's function in hematopoiesis.

Figure 24: Effect of over expression of MLL-Nwt and mutants on the colony formation potential of hematopoietic progenitors



Ckit+ve selection and transduction as explained in Figure. Transduced cells were selected with G418 on a methyl cellulose plate containing IL-3, IL-6, SCF, and GM-CSF for 7 days. Colonies were counted on Day7. n=3, average of 3 independent

Figure 25: Effect of MLL-N wt and mutants over expression on colony formation of MLL-AF9 transformed hematopoietic progenitors



C-kit+ve hematopoietic progenitors were isolated from mouse bone marrow, transduced with MSCV puro MLL-AF9 retoviruses, selected for a week in puromycin and grown in RPMI containing IL-3, IL-6, SCF and GM-CSF . Further, cells were transduced with MSCV neo (Vector), MLL-N wt, M1585A or M1606D retroviruses and plated on methyl cellulose containing G418, IL-3, IL-6, SCF, and GM-CSF. Colonies were counted on Day7. The cells were recovered and re-plated on methyl cellulose. The figure above is from week3 re-plating. Representative colony images are shown. n=2, average of two independent experiments, error bar represents S.D.

The results from this study revealed that MLL PHD3 binding to H3K4Me3 contributes to transcriptional activation of genes that regulate hematopoiesis. Recent studies showed that the MLL wt allele is required for MLL-fusion mediated leukemogenesis [20,40]. The MLL-fusion protein lacks the 3rd PHD finger. But the MLL wt allele has the PHD3 intact. To know if MLL PHD3 (present in MLL wt) binding to H3K4Me3 contributes to MLL fusion mediated leukemogenesis, the effect of overexpression of MLL-N wt and the mutants on colony formation by MLL-AF9 transformed hematopoietic progenitors was tested (Figure 25). Over expression of MLL-N or the mutants did not affect the colony formation of MLL-AF9 transformed hematopoietic progenitors in the assay conditions tested.

CHAPTER 5

DISCUSSION

Many chromatin associated proteins have tandem PHD fingers that mediate DNA binding, histone binding or serve as domains for other protein-protein interactions. MLL has 3 tandem PHD fingers. More than a decade ago, studies from our lab identified Cyp33 as a ligand for MLL PHD3. Cyp33 over expression resulted in transcriptional repression of MLL target genes [23,57]. Thus Cyp33 over expression switches MLL from a transcriptional activator to a repressor. Recently we, and two other groups reported that the MLL PHD3 binds to trimethylated H3K4 (H3K4Me3) [26,29,30]. Since H3K4Me3 is a transcriptional activation mark, we hypothesized that binding of MLL PHD3 to H3K4Me3 protects it from histone demethylases which would result in transcriptional activation of MLL target genes. Thus the PHD3, a small 60 amino acids long region of MLL can potentially act as a molecular switch in regulation of transcription of its target genes.

In vitro binding studies performed by our collaborator Dr. Bushweller, showed that when Cyp33 is in excess, it reduces the affinity of the MLL PHD3 for H3K4Me3 [26]. To determine if this is true *in vivo*, I sought to test the effect of Cyp33 over expression in 293T cells on H3K4Me3 methylation and binding of histone demethylases to MLL target gene promoters. Previous studies employed a semi- quantitative approach to measure

gene expression [23], while this study employs qRT-PCR, a quantitative method to measure gene expression. Since a different technique was used than in previous studies, before doing ChIP for H3K4Me3, the effect of Cyp33 over expression on transcription of MLL target genes was first confirmed. We confirmed that Cyp33 over expression resulted in a decrease in transcription of MLL target genes. However, the effect of Cyp33 over expression on transcription is not exclusive of MLL target genes. Except for *β2M*, other house-keeping genes, which are not MLL target genes, also showed transcriptional repression upon Cyp33 over expression. Recently, Steven Poppen from our lab has shown that both over expression and knock down of Cyp33 resulted in slow growth of cells in culture. This suggests that altering the levels of Cyp33 in the cell may have an effect on general metabolism of the cell. Besides, binding to MLL PHD3, Cyp33 is known to be present in the spliceosome complex [107]. Hence the effect of Cyp33 over expression on transcription of house-keeping genes may be MLL dependent or MLL independent.

Having confirmed that Cyp33 over expression resulted in decreased expression of MLL target genes, I wanted to determine the effect of excess of Cyp33 on H3 density, transcriptional activation and repressive histone marks. H3 density reflects the closed or open status of the chromatin for access to the RNA POLII transcriptional machinery. Cyp33 over expression resulted in increased H3density at MLL target gene promoters, which indicates a closed chromatin state and is consistent with the transcriptional repression of target genes. One of the possible mechanisms of Cyp33 mediated transcriptional repression of MLL target genes, is through increased recruitment of co-repressors like HDAC1 and Bmi-1 to the repression domain of MLL [27]. Results from this

study show that Cyp33 over expression results in a decrease in H3 acetylation marks which can be explained by the increased recruitment of HDAC to the repression domain of MLL.

Like H3Ac, increased H3K4Me3 at the promoter region is linked to transcriptional activation. As mentioned earlier, from *in vitro* binding studies, it is known that when Cyp33 is in excess, it reduces the affinity of the MLL PHD3 for the H3K4Me3 histone mark [26]. *In vivo* the scenario is more complicated. The 3rd PHD finger of the JARID1A and JARID1B histone demethylases is also known to bind H3K4Me3 [35,101]. The reduced affinity of the MLL PHD3 for H3K4Me3 after binding of Cyp33 may allow histone demethylases or other competing ligands to bind this histone mark. According to this hypothesis, Cyp33 over expression should result in increased recruitment of JARID1A and JARID1B histone demethylases and consequent decrease in H3K4Me3 marks. Results from our ChIP experiment showed that Cyp33 over expression resulted in increased recruitment of JARID1A and JARDI1B and a corresponding decrease in H3K4Me3 marks at MLL target gene promoters which is in agreement with our hypothesis.

Concordant with its effects on the transcription of house-keeping genes, Cyp33 over expression also resulted in increased recruitment of JARID1B at house-keeping gene promoters. However, Cyp33 over expression did not affect the transcription of *JARID1A* and *JARID1B* nor did it increase the recruitment of *JARID1B* at the *JARID1B* promoter. This suggests that the effect of Cyp33 over expression on transcription may not be a global phenomenon and occurs at a subset of genes. Knock down of Cyp33 in an experiment performed in collaboration with Steven Poppen showed a significant decrease

in JARID1B recruitment at MLL target gene promoters which is the opposite of what we observe when over expressing Cyp33.

Cyp33 when present in excess, besides exposing the methyl marks to JARID1 histone demethylases, can also increase the recruitment of JARID1 histone demethylases at MLL target genes by other mechanisms. When compared to JARID1A, JARID1B recruitment is more pronounced upon Cyp33 over expression. JARID1B directly interacts with Class I and Class II HDACs and with hPC2 [105,108], but there are no studies so far showing a direct interaction of JARID1A with HDAC and hPC2. As mentioned earlier, Cyp33 over expression, augments the recruitment of transcriptional co-repressors like, HDAC1 and Bmi1 to the repression domain of MLL [27]. It is certainly a possibility that along with other co-repressors Cyp33 may also bring in JARID1B to the repression domain of MLL.

Another mechanism by which Cyp33 can increase the recruitment of JARID1A and JARID1B at MLL target promoters is through directly binding to JARID1A and JARID1B. Besides the PHD finger of MLL, PHD fingers present in other proteins are not known to bind Cyp33. JARID1A and JARID1B have 3 PHD fingers. Results from this study shows that Cyp33 does not co-immunoprecipitate with JARID1B.

Cyp33 is a prolyl isomerase, inhibiting the prolyl isomerase activity by Cyclosporin A, deleting the Cyclophilin domain that encodes for the PPIase, or over expressing a prolyl isomerase deficient mutant of Cyp33, does not result in transcriptional repression of MLL target genes [23,26]. This shows that the PPIase activity of Cyp33 is required for the transcriptional repression of MLL target genes. Additionally,

the PPIase activity is required for the recruitment of transcriptional repressors like HDAC1 and Bmi1 to the repression domain of MLL. Like HDAC1 and Bmi-1, JARID1A and JARID1B are transcriptional repressors and are present in complex with HDAC. The results from this study show that recruitment of JARID1A and JARID1B is PPIase dependent. Furthermore, over expression of the PPIase mutant results in decrease in H3K4Me3 levels and H3K27Ac levels at MLL target gene promoters. Though we were not able to pin point the exact mechanism of Cyp33 mediated increased recruitment of JARID1A and JARID1B to the chromatin, we have identified the requirement of PPIase activity of Cyp33 for JARID1A and JARID1B recruitment at MLL target gene promoters.

Since Cyp33 over expression studies suggested a role for JARID1B in transcriptional regulation of MLL target genes, shRNA knock down of JARID1B was performed in 293T cells. Knock down of JARID1B resulted in increased expression of *HOXA9* and *CDKN1B,* indicating that JARID1B is a transcriptional repressor of a sub-set of MLL target genes. Only 50% reduction in JARID1B expression was achieved by our knock down. If the efficiency of the knock down was improved, there may be an effect on the expression of other MLL target genes like *HOXC8* and *MYC*. JARID1A and JARID1B are redundant in H3K4Me3 demethylation [31]. The loss of JARID1B can be functionally compensated by JARID1A. Hence a combination of JARID1A and JARID1B knock down may have a greater effect on MLL target gene expression.

Cyp33 over expression studies and JARID1B knock down studies suggest that JARID1B is a downstream effector of Cyp33. If JARID1B is a downstream effector of Cyp33-mediated transcriptional regulation of MLL target genes, knock down of

JARID1B should overcome the transcriptional repression resulting from Cyp33 over expression. But our results show that Cyp33 over expression, even after the knock down of JARID1B, resulted in transcription repression of MLL target genes. From this experiment we cannot conclude if JARID1B is a downstream effector of MLL target genes. The caveat of this experiment is that JARID1B may not be the only downstream effector of Cyp33. As mentioned earlier, there is redundancy in the function of JARID1A and JARID1B, and even the knockdown of JARID1B was incomplete. The ideal experiment to do would be to over express Cyp33 in a cell line that is null for JARID1A and JARID1B (MEF JARID1A-/- , JARID1B -/-). Since Cyp33 can also operate through other transcriptional co-repressors like HDAC1at MLL target gene promoters, it may be difficult to prove JARID1B as one of the downstream effectors of Cyp33.

The results from the first aim of this study demonstrated the effects of excess of Cyp33 on MLL target gene expression and more specifically on H3K4 trimetylation. The biological significance of MLL PHD3 binding to H3K4Me3 and Cyp33 is not known. Chang *et al*., have shown that the abrogation of H3K4Me3 binding to MLL PHD3 resulted in decreased transcriptional activation of MLL target genes [30]. MLL is a transcriptional regulator of genes involved in hematopoiesis. To determine the biological significance of MLL PHD3 binding to H3K4Me3 versus Cyp33, the effect of abrogation of binding was studied using an *in vitro* colony formation assay of hematopoietic progenitors. To do so, retroviral transduction of wild type MLL-N, and MLL-N mutants, in mouse hematopoietic progenitors was performed. Due to its molecular weight, it is difficult to package full length *MLL* cDNA in a retro viral vector. MLL is proteolytically cleaved into MLL-N and MLL-C by TASPASE-1. The PHD3 in MLL-N is the reader of

H3K4Me3 and the SET domain in MLL-C is the writer for H3K4Me3. Thus by making mutations in MLL-N fragment, instead of full length MLL, will allow us to precisely study the effects of abrogation of the reader function without over expressing the writer. Furthermore, it will also help to overcome the practical difficulties of using a very big construct for retro viral packaging.

Similar to the PHD fingers of BPTF and ING family members that bind H3K4Me3, the conserved aromatic residues in MLL PHD3 form an aromatic cage which is the binding pocket for H3K4Me3. Chang et al. have shown that this mutation resulted in decreased binding of H3K4Me3 without affect the protein stability [30]. Hence this was chosen as a candidate mutation that would abrogate the binding of MLL PHD3 to H3K4Me3. The H3K4Me3 binding site and Cyp33 binding site in MLL PHD3 are distinct, but very close to each other. Hence abrogating one binding may affect the other. Wang et al. showed that the M1606D mutation abolished the binding of MLL PHD3 to Cyp33. However the effect of this mutation on H3K4Me3 binding is not known. Before doing any hematopoietic progenitor colony formation assays, the binding of MLL-wt, M1585A, and M1606D mutants to H3K4Me3 *in vivo* was characterized. The results from nucleosome Co-IP showed that the MLL-N wt and M1606D mutant co-immunoprecipitated with nucleosomes trimethylated on H3K4 whereas the M1585A mutant MLL-N did not.

To date, there are no studies on the effect of overexpression of MLL-N wt on target gene expression. To study this and to know the effect of MLL-N mutants on target gene expression, over expression of MLL-N wt and mutants was induced in MEFs using retroviral expression constructs. Hsieh *et al*., have shown that MLL-C is required for the

stability of MLL-N [74]. Based on this finding, over expressing just the MLL-N fragment should result in degradation of excess of MLL-N. Therefore over expression of MLL-N wt should not have much effect on target gene expression. Surprisingly, MLL-N wt over expression resulted in increased MLL target gene expression both in MEF and in hematopoietic progenitors. Yokoyoma *et al*., have shown that free MLL-C in the nucleus is targeted for proteosomal degradation in the cytoplasm [38]. If the endogenous levels of MLL-C are not limiting, and over expression of MLL-N enhances the MLL-N-C holo-complex formation in the nucleus, then upon over expression of MLL-N there should be increased levels of MLL-C in the nucleus. However the results show that MLL-N over expression did not influence the protein levels in the nucleus, but resulted in increased recruitment of MLL-C and a corresponding increase in H3K4Me3 marks at MLL target gene promoters compared to the vector control.

Unlike MLL-N wt over expression; the M1585A mutation did not result in transcriptional activation of MLL target genes. This proves that MLL PHD3 binding to H3K4Me3 contributes to the transcriptional activation of MLL target genes. M1585A over expression did not result in increased recruitment of MLL-C or an increase of H3K4Me3 marks. This suggests that MLL PHD3 binding to H3K4Me3 marks recruits more MLL-C and helps in methylation of adjacent nucleosomes leading to transcriptional activation of MLL target genes.

The results from Cyp33 over expression experiments performed in the first aim of this study showed a Cyp33 mediated increase in JARID1A/B demethylases at MLL target genes and decrease in H3K4Me3 marks at MLL target gene promoters. This result suggested a role for MLL PHD3 in protecting H3K4Me3 from demethylation. If that is

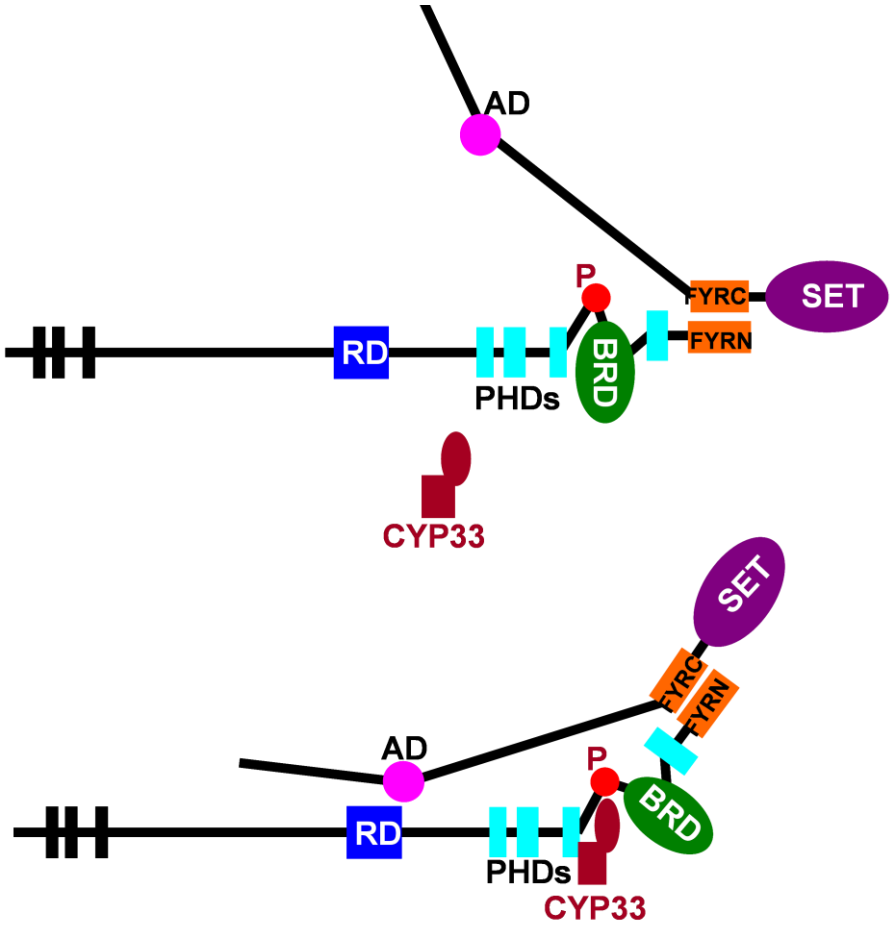
true, then upon abrogation of MLL PHD3 binding to H3K4Me3, there should be increased recruitment of JARID1 histone demethylases and a corresponding decrease in H3K4Me3. However, mutating the H3K4Me3 binding site in MLL PHD3 (M1585A) did not result in increased recruitment of JARID1B at MLL target gene promoters, while showing decreased recruitment of MLLC and H3K4 trimethylation. This result suggest that in MEFs (*Mll*+/+), MLL PHD3 binding to H3K4Me3 is mainly required for propagation of H3K4 methyl marks to adjacent nucleosomes and not for the protection of H3K4Me3 from JARID1B histone demethylase. Additionally, this result indicates that there may be alternative mechanisms that contribute to Cyp33 mediated increased recruitment of JARID1B at MLL target gene promoters other than Cyp33 reducing the protection of H3K4Me3 by MLL PHD3.

Cyp33 is a transcriptional repressor of MLL target genes; therefore, by abrogating Cyp33 binding to MLL PHD3, we expected a greater increase in MLL target genes expression than MLL-N wt over expression. Over expression of M1606D mutant (that does not bind Cyp33) resulted in increased expression of MLL target genes, but the effect was similar to that of MLL-N wt over expression. This may be because the repressive effects of endogenous levels of Cyp33 in the cell may be too small in comparison with the transactivation effect of overexpression of MLL-N to be noticed in this experiment. Additionally, over expression of the M1606D mutant resulted in decreased recruitment of MLL-C and JARID1B when compared to MLL-N wt and M1585A at the *Hoxa9* promoter. The results from the first aim of the study suggested that JARID1B is one of the downstream effectors of Cyp33 mediated transcriptional repression of MLL target genes. In agreement with this result, mutating the Cyp33 binding site (M1606D) in the

MLL PHD3 resulted in decreased recruitment of JARID1B at the *HOXA9* promoter. Surprisingly, this mutation also resulted in decreased recruitment of MLL-C at the *Hoxa9* promoter. Despite the decreased recruitment of MLL-C, the increase in expression of MLL target genes upon over expression of M1606D may be explained by relief of transcriptional repression by JARID1B. The reason for the over expression of M1606D mutation resulting in decreased recruitment of MLL-C at the *HOXA9* promoter is not clear. Yokoyama *et al*., have shown that PHD1, PHD4 and FYRN and FYRC domains are required for interaction of MLL-N and MLL-C [38]. Since the MLL PHD 1, 2 and 3 are tandem PHD fingers; the methionine to aspartate mutation in PHD3 may have an effect on structure and function of the 3 PHD cassette unit and thus may interfere with MLL-C recruitment by PHD1. Additionally, the targeting of proline 1629 by Cyp33 from cis-to trans results in a drastic change in the relative positions of PHD1 and PHD4 which are located at the opposite ends of this prolyl bond, which may contribute to the recruitment of MLL-C by PHD-1 and PHD-4 (Figure 26) [29,38].

The *Hoxa9* and *Meis1* genes are expressed at higher levels in hematopoietic progenitors and are required to keep the hematopoietic progenitors in an undifferentiated state [106]. Over expression of MLL-N wt and M1606D in hematopoietic progenitors resulted in increased expression of MLL target genes like *Hoxa9* and *Meis1* and also increased the colony formation of hematopoietic progenitors when compared to M1585A (which does not bind H3K4Me3). This proves that MLL PHD3 binding to H3K4Me3 contributes to transcriptional activation of MLL target genes allowing the hematopoietic progenitors to remain in undifferentiated state and form colonies.

Figure 26: Putative model for the effect of Cyp33 binding to MLL PHD3 on recruitment of MLL-C



The prolyl-isomerase activity of Cyp33 targets P1629 in the linker region between the PHD3 and the bromo-domain of MLL. This cis- trans peptidyl prolyl isomerization results in drastic change in conformation and the relative positions of PHD1 and PHD4 which are located at the opposite ends of this prolyl bond and allows Cyp33 to bind to PHD3[29] . In addition to FYRN and FYRC domains of MLL; the PHD1 and PHD4 are also required for mediating the interaction between MLL-N and MLL-C[38]. Abrogation of MLL PHD3 binding to Cyp33 PHD3 (M1606D) may have an effect on the structure and the function of the 3 PHD cassette unit and thus may interfere with MLL-C recruitment by PHD1.

Cdkn1b is a transcriptional target of MLL. Xia *et al*., have shown that the transcriptional outcome of MLL overexpression depends on the cell type [11]. Concordant with Xia et al. results, MLLN over expression showed cell-type specific transcriptional outcome of *Cdkn1b*. As mentioned earlier, MLL-N wt over expression resulted in transcriptional activation of *Cdkn1b* in MEFs, whereas, there was no effect in hematopoietic progenitors. In contrast to MLL-N wt, over expression of M1585A in hematopoietic progenitors resulted in decreased expression of self-renewal genes like *Hoxa9* and *Meis1*, but showed increased expression of cell cycle inhibitors like *Cdkn1b* and *Cdkn1a*. Furthermore, over expression of M1585A in hematopoietic progenitors resulted in drastic reduction of colony formation. Fast proliferating cells like hematopoietic progenitors have lower levels of *Cdkn1b* whereas growth arrested cells like terminally differentiated cells have higher levels of *Cdkn1b* expression [109]. *Cdkn1a* expression increases upon induction of differentiation [109]. Taken together, the decreased expression of self-renewal genes, increased expression of cell cycle regulators and decreased colony formation suggesting that M1585A over expression may inhibit the proliferation and/or induce differentiation of hematopoietic progenitors.

The wild type *MLL* allele is required for MLL-AF9 induced leukemogenesis and is also required for the survival and maintenance of MLL-AF9 transformed leukemic cells. The wild type *MLL* is recruited to *HOXA9* and *Meis1* promoters and is required for trimethylation of H3K4 at *HOXA9* and *Meis1* promoters, thus contributing to their transcriptional activation in MLL-AF9 transformed leukemia cells [20]. MLL fusions lack PHD fingers. We hypothesized that MLL PHD3 (encoded in the wild type MLL allele) binding to H3K4Me3 contributes to the transcriptional activation of MLL target genes in

MLL-AF9 transformed cells. The results from this experiment showed that MLL-N or mutants over expression did not have an effect on colony formation of MLL-AF9 transformed hematopoietic progenitors. The MLL fusion protein constitutively activates the transcription of *Hoxa9* and *Meis1*. Most common fusion partners of MLL; AF9, AF4, ELL and ENL are present in super elongation complex (SEC) along with transcription elongation factor like P-TEFB. The SEC is implicated in transcriptional activation of *Hox* genes in MLL rearranged leukemia [110]. The common fusion partners like AF-9, AF-10, and ENL are found in complex with DOT1L, a H3K79 methyl transferase. The aberrant recruitment of DOT1L to MLL target promoters constitutively activates the transcription of *Hox*a9 [111–114]. Hence the effects of MLL-N wt or mutants may be masked by the very high levels of over expression of MLL-AF9 in mouse hematopoietic progenitors. In the future, to determine the role of MLL PHD3 binding to H3K4Me3, we may have to partially inhibit DOT1L in MLL-AF9 transduced hematopoietic progenitors and then study the effects of abrogation of MLL PHD3 on leukemogenesis.

The results from the experiments performed in mouse hematopoietic progenitors showed that MLL PHD3 binding to H3K4Me3 contributes to the transcriptional activation of MLL target genes in hematopoietic progenitors and is required for hematopoietic progenitor proliferation. Thus we have identified a novel biological function for MLL PHD3 binding to H3K4Me3, this opens up many avenues for further studies. Chromosomal amplification of *MLL* is frequently seen in MDS, AML, ALL, and specifically in therapy related leukemia and is associated with poor prognosis [115–118]. Patients with *MLL* amplification over express the wild type allele and show increased expression of *HOXA9* and Meis1. Hence disrupting the interaction between MLL PHD3

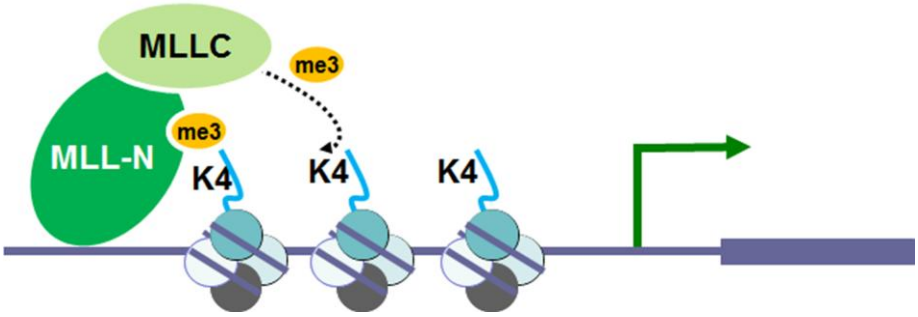
and H3K4ME3 in *MLL* amplified leukemia may be a potential treatment option [115]. *MLL* is haplo-insufficient, given the importance of MLL in fetal and adult differentiation, development and cell cycle regulation, a heterozygous mutation that disrupts the binding of MLL PHD3 to H3K4Me3 may prevent MLL from performing its functions resulting in developmental disorders or diseases. Due to the enormous size of the MLL gene, the mutation studies on MLL are limited. However with the advent of next generation sequencing, it will be interesting to determine if the mutation of MLL PHD3 binding to H3K4Me3 occurs frequently in any developmental disorders.

In summary, over expression of Cyp33 caused increased H3 density, decreased H3Ac, increased recruitment of JARID1 histone demethylases, and decreased H3K4Me3 at MLL target gene promoters. As a consequence, Cyp33 over expression resulted in transcription repression of MLL target genes. Abrogation of binding of H3K4Me3 in MLL PHD3 resulted in decreased recruitment of MLL-C, decreased in H3K4Me3 levels, decreased MLL target gene expression and decreased expansion of hematopoietic progenitors when compared to MLL-N wt, demonstrating that MLL PHD3 binding to H3K4Me3 contributes to the transcriptional activation of MLL target genes.

Collectively, the results from this study points us to a model wherein MLL PHD3 binding to H3K4Me3 helps MLL-C to trimethylate H3K4 on adjacent nucleosomes, contributing to transcriptional activation of MLL target genes, as shown in the model (Figure 27). Whereas, excess of Cyp33 results in increased recruitment of JARID1A/B at MLL target gene promoters in a PPIase- dependent manner resulting in active demethylation of H3K4Me3, resulting in target gene repression (Figure 28). Thus MLL

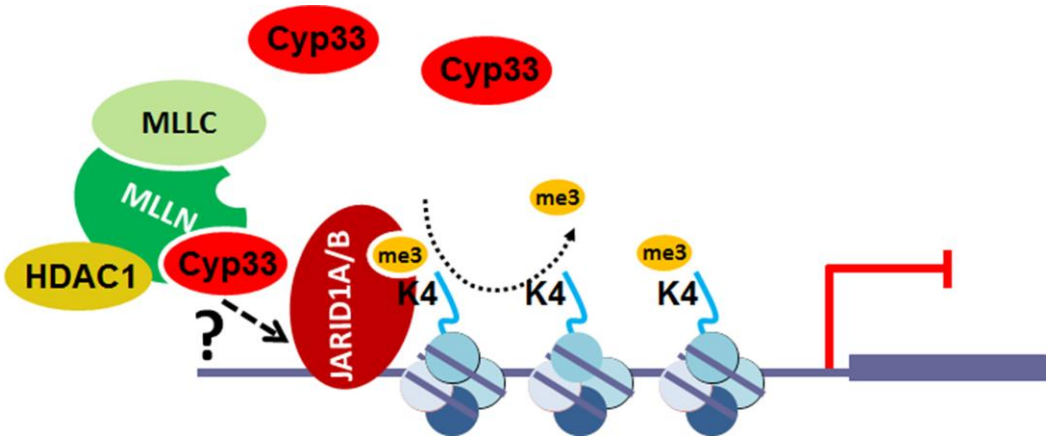
PHD3 functions as a part of a molecular switch in transcriptional regulation of MLL target genes.

Figure 27: Model for the function of the MLL PHD3-H3K4Me3 interaction



MLL is recruited to its target promoters by several mechanisms. The binding of MLL PHD3 to H3K4Me3 may be one of the mechanisms. Thus the MLL PHD3 may help target the SET domain of MLL-C to trimethylate adjacent histones. This maintains the H3K4Me3 marks at the target promoters leading to sustained expression of *Hoxa9* and *Meis1* in hematopoietic progenitors.

Figure 28: Model for effect of excess of Cyp33 on trimethylation of H3K4 at MLL target gene promoters



Excess of Cyp33 results in increased recruitment of JARID1 A/B histone demethylases to MLL1 target gene promoters, resulting in decrease in H3K4Me3 marks at MLL1 target gene promoters leading to transcriptional repression of MLL1 target genes. The mechanism by which Cyp33 binding to MLL stimulates JARID1A/B recruitment to chromatin is not known. However, the PPIase activity of Cyp33 is required for the recruitment of JARID1B to MLL target gene promoters. JARID1B binds DNA in a sequence specific manner, and binds H3K4Me3.

REFERENCES

1. Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E. & Korsmeyer, S. J. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 10632–6 (1998).

2. Hughes, C. M. *et al.* Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. *Molecular cell* **13**, 587–97 (2004).

3. Milne, T. A. *et al.* MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular Cell* **10**, 1107–1117 (2002).

4. Shilatifard, A. The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. *Annual Review of Biochemistry* **81**, 65–95 (2012).

5. Nakamura, T. *et al.* ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Molecular Cell* **10**, 1119–1128 (2002).

6. Hess, J. L., Yu, B. D., Li, B., Hanson, R. & Korsmeyer, S. J. Defects in yolk sac hematopoiesis in Mll-null embryos. *Blood* **90**, 1799–1806 (1997).

7. Gan, T., Jude, C. D., Zaffuto, K. & Ernst, P. Developmentally induced Mll1 loss reveals defects in postnatal haematopoiesis. *Leukemia official journal of the Leukemia Society of America Leukemia Research Fund UK* **24**, 1732–1741 (2010).

8. Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. & Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505–508 (1995).

9. Tyagi, S., Chabes, A. L., Wysocka, J. & Herr, W. E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Molecular Cell* **27**, 107–119 (2007).

10. Milne, T. A. *et al.* Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 749–754 (2005).

11. Xia, Z.-B. *et al.* The MLL fusion gene, MLL-AF4, regulates cyclin-dependent kinase inhibitor CDKN1B (p27kip1) expression. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14028–14033 (2005).

12. Ziemin-van Der Poel, S. *et al.* Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 10735–10739 (1991).

13. Thirman, M. J. *et al.* Rearrangement of the MLL gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations. *The New England Journal of Medicine* **329**, 909–914 (1993).

14. Thorsteinsdottir, U., Kroon, E., Jerome, L., Blasi, F. & Sauvageau, G. Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. *Molecular and Cellular Biology* **21**, 224–234 (2001).

15. Zeisig, B. B. *et al.* Hoxa9 and Meis1 Are Key Targets for MLL-ENL-Mediated Cellular Immortalization. *Molecular and Cellular Biology* **24**, 617–628 (2004).

16. Wang, Q.-F. *et al.* MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood* **117**, 6046–6047 (2011).

17. Ayton, P. M. & Cleary, M. L. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes & Development* **17**, 2298–2307 (2003).

18. Liedtke, M. & Cleary, M. L. Therapeutic targeting of MLL. *Blood* **113**, 6061–6068 (2009).

19. Bernt, K. M. & Armstrong, S. A. Targeting epigenetic programs in MLL-rearranged leukemias. *Hematology the Education Program of the American Society of Hematology American Society of Hematology Education Program* **2011**, 354–60 (2011).

20. Thiel, A. T. *et al.* MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. *Cancer Cell* **17**, 148–159 (2010).

21. Zeleznik-Le, N. J., Harden, A. M. & Rowley, J. D. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 10610–10614 (1994).

22. Ayton, P. M. & Cleary, M. L. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene* **20**, 5695–5707 (2001).

23. Fair, K. *et al.* Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Molecular and Cellular Biology* **21**, 3589–3597 (2001).

24. Wang, Y. *et al.* Human CyP33 binds specifically to mRNA and binding stimulates PPIase activity of hCyP33. *FEBS Letters* **582**, 835–839 (2008).

25. Mi, H., Kops, O., Zimmermann, E., Jäschke, A. & Tropschug, M. A nuclear RNA-binding cyclophilin in human T cells. *FEBS Letters* **398**, 201–205 (1996).

26. Park, S. *et al.* The PHD3 domain of MLL acts as a CYP33-regulated switch between MLL-mediated activation and repression . *Biochemistry* **49**, 6576–86 (2010).

27. Xia, Z.-B., Anderson, M., Diaz, M. O. & Zeleznik-Le, N. J. MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8342–7 (2003).

28. Chen, J. *et al.* Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Research* **68**, 6199–6207 (2008).

29. Wang, Z. *et al.* Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell* **141**, 1183–94 (2010).

30. Chang, P.-Y. *et al.* Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. *Journal of molecular biology* **400**, 137–44 (2010).

31. Klose, R. J. *et al.* The retinoblastoma binding protein RBP2 is an H3K4 demethylase. *Cell* **128**, 889–900 (2007).

32. Yamane, K. *et al.* PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Molecular Cell* **25**, 801–812 (2007).

33. Christensen, J. *et al.* RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell* **128**, 1063–76 (2007).

34. Xiang, Y. *et al.* JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 19226–31 (2007).

35. Musselman, C. A. & Kutateladze, T. G. Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Research* **39**, 1–11 (2011).

36. Hsieh, J. J., Ernst, P., Erdjument-bromage, H., Tempst, P. & Korsmeyer, S. J. Proteolytic Cleavage of MLL Generates a Complex of N- and C-Terminal Fragments That Confers Protein Stability and Subnuclear Localization. *Society* **23**, 186–194 (2003).

37. Liu, H., Cheng, E. H.-Y. & Hsieh, J. J.-D. Bimodal degradation of MLL by SCFSkp2 and APCCdc20 assures cell cycle execution: a critical regulatory circuit lost in leukemogenic MLL fusions. *Genes & Development* **21**, 2385–2398 (2007).

38. Yokoyama, A. *et al.* Proteolytically cleaved MLL subunits are susceptible to distinct degradation pathways. *Journal of Cell Science* **124**, 2208–2219 (2011).

39. Ernst, P., Wang, J., Huang, M., Goodman, R. H. & Korsmeyer, S. J. MLL and CREB Bind Cooperatively to the Nuclear Coactivator CREB-Binding Protein. *Molecular and Cellular Biology* **21**, 2249–2258 (2001).

40. Milne, T. a *et al.* Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Molecular cell* **38**, 853–63 (2010).

41. Meyer, C. *et al.* New insights to the MLL recombinome of acute leukemias. *Leukemia official journal of the Leukemia Society of America Leukemia Research Fund UK* **23**, 1490–1499 (2009).

42. Muntean, A. G. & Hess, J. L. The Pathogenesis of Mixed-Lineage Leukemia. *Annual Review of Pathology* **7**, 283–301 (2011).

43. Dolan, M., McGlennen, R. C. & Hirsch, B. MLL amplification in myeloid malignancies: clinical, molecular, and cytogenetic findings. *Cancer Genetics and Cytogenetics* **134**, 93–101 (2002).

44. Caligiuri, M. A. *et al.* Rearrangement of ALL1 (MLL) in acute myeloid leukemia with normal cytogenetics. *Cancer Research* **58**, 55–59 (1998).

45. Grembecka, J., Belcher, A. M., Hartley, T. & Cierpicki, T. Molecular basis of the mixed lineage leukemia-menin interaction: implications for targeting mixed lineage leukemias. *The Journal of Biological Chemistry* **285**, 40690–40698 (2010).

46. Yokoyama, A. *et al.* The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123**, 207–218 (2005).

47. Caslini, C. *et al.* Interaction of MLL amino terminal sequences with menin is required for transformation. *Cancer Research* **67**, 7275–7283 (2007).

48. Yokoyama, A. & Cleary, M. L. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* **14**, 36–46 (2008).

49. Jin, S. *et al.* c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. *Journal of Clinical Investigation* **120**, 593–606 (2010).

50. Adler, H. T., Nallaseth, F. S., Walter, G. & Tkachuk, D. C. HRX leukemic fusion proteins form a heterocomplex with the leukemia-associated protein SET and protein phosphatase 2A. *The Journal of Biological Chemistry* **272**, 28407–28414 (1997).

51. Adler, H. T. *et al.* Leukemic HRX fusion proteins inhibit GADD34-induced apoptosis and associate with the GADD34 and hSNF5/INI1 proteins. *Molecular and Cellular Biology* **19**, 7050–7060 (1999).

52. Yano, T. *et al.* Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 7286–7291 (1997).

53. Cierpicki, T. *et al.* Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. *Nature Structural & Molecular Biology* **17**, 62–68 (2010).

54. Erfurth, F. E. *et al.* MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7517–7522 (2008).

55. Muntean, A. G. *et al.* The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell* **17**, 609–621 (2010).

56. Wang, J., Muntean, A. G., Wu, L. & Hess, J. L. A subset of mixed lineage leukemia protein has PHD-domain mediated E3 ligase activity. *Journal of Biological Chemistry* (2012).

57. Anderson, M. *et al.* A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in Drosophila and human cells. *Development Genes and Evolution* **212**, 107–113 (2002).

58. Dou, Y. *et al.* Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873–885 (2005).

59. Yokoyama, A. *et al.* Leukemia Proto-Oncoprotein MLL Forms a SET1-Like Histone Methyltransferase Complex with Menin To Regulate Hox Gene Expression. *Molecular and Cellular Biology* **24**, 5639–5649 (2004).

60. Rozovskaia, T. *et al.* Self-association of the SET domains of human ALL-1 and of Drosophila TRITHORAX and ASH1 proteins. *Oncogene* **19**, 351–357 (2000).

61. Krajewski, W. A., Nakamura, T., Mazo, A. & Canaani, E. A Motif within SET-Domain Proteins Binds Single-Stranded Nucleic Acids and Transcribed and Supercoiled DNAs and Can Interfere with Assembly of Nucleosomes. *Molecular and Cellular Biology* **25**, 1891–1899 (2005).

62. Liu, H. *et al.* Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint. *Nature* **467**, 343–346 (2010).

63. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. & Cavalli, G. Genome regulation by polycomb and trithorax proteins. *Cell* **128**, 735–45 (2007).

64. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual Review of Genetics* **38**, 413–43 (2004).

65. Breen, T. R. Mutant alleles of the Drosophila trithorax gene produce common and unusual homeotic and other developmental phenotypes. *Genetics* **152**, 319–344 (1999).

66. Terranova, R., Agherbi, H., Boned, A., Meresse, S. & Djabali, M. Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6629–6634 (2006).

67. McMahon, K. a *et al.* Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell stem cell* **1**, 338–45 (2007).

68. Guenther, M. G. *et al.* Global and Hox-specific roles for the MLL1 methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8603–8 (2005).

69. Milne, T. A. *et al.* MLL associates specifically with a subset of transcriptionally active target genes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14765–14770 (2005).

70. Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Molecular and cellular biology* **29**, 6074–85 (2009).

71. Krumlauf, R. Hox genes in vertebrate development. *Cell* **78**, 191–201 (1994).

72. Argiropoulos, B. & Humphries, R. K. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* **26**, 6766–76 (2007).

73. Huang, G. *et al.* PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nature Genetics* **40**, 51–60 (2008).

74. Takeda, S. *et al.* Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes & development* **20**, 2397–409 (2006).

75. R.D Kornberg Structure of Chromatin. *Annual Review of Biochemistry* **46**, 931–954 (1977).

76. Finch, J. T. & Klug, A. Solenoidal model for superstructure in chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 1897–1901 (1976).

77. McGhee, J. D., Felsenfeld, G. & Eisenberg, H. Nucleosome structure and conformational changes. *Biophysical Journal* **32**, 261–270 (1980).

78. Dinant, C., Houtsmuller, A. B. & Vermeulen, W. Chromatin structure and DNA damage repair. *Epigenetics chromatin* **1**, 9 (2008).

79. Paranjape, S. M., Kamakaka, R. T. & Kadonaga, J. T. Role of chromatin structure in the regulation of transcription by RNA polymerase II. *Current Opinion in Cell Biology* **58**, 417–423 (1993).

80. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).

81. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–80 (2001).

82. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349–352 (1997).

83. Lee, K. K. & Workman, J. L. Histone acetyltransferase complexes: one size doesn't fit all. *Nature Reviews Molecular Cell Biology* **8**, 284–295 (2007).

84. Hildmann, C., Riester, D. & Schwienhorst, A. Histone deacetylases--an important class of cellular regulators with a variety of functions. *Applied Microbiology and Biotechnology* **75**, 487–497 (2007).

85. Yun, M., Wu, J., Workman, J. L. & Li, B. Readers of histone modifications. *Cell Research* **21**, 564–578 (2011).

86. Blair, L. P., Cao, J., Zou, M. R., Sayegh, J. & Yan, Q. Epigenetic Regulation by Lysine Demethylase 5 (KDM5) Enzymes in Cancer. *Cancers* **3**, 1383–1404 (2011).

87. Hom, R. a *et al.* Molecular mechanism of MLL PHD3 and RNA recognition by the Cyp33 RRM domain. *Journal of molecular biology* **400**, 145–54 (2010).

88. Takahashi, N., Hayano, T. & Suzuki, M. Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin. *Nature* **337**, 473–475 (1989).

89. Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. & Schmid, F. X. Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. *Nature* **337**, 476–478 (1989).

90. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436–439 (2005).

91. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).

92. Secombe, J. & Eisenman, R. N. The function and regulation of the JARID1 family of histone H3 lysine 4 demethylases: the Myc connection. *Cell cycle Georgetown Tex* **6**, 1324–1328 (2007).

93. Cloos, P. A. C., Christensen, J., Agger, K. & Helin, K. cellular differentiation and disease Erasing the methyl mark : histone demethylases at the center of cellular differentiation and disease. 1115–1140 (2008).doi:10.1101/gad.1652908

94. Schindler, U., Beckmann, H. & Cashmore, A. R. HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *The Plant Journal* **4**, 137–150 (1993).

95. Li, Y. & Li, H. Many keys to push : diversifying the " readership " of plant homeodomain fingers Recognition of Histone Tail by Single. *Acta Biochim Biophys Sin* **4**, 28–39 (2012).

96. Bienz, M. The PHD finger, a nuclear protein-interaction domain. *Trends in biochemical sciences* **31**, 35–40 (2006).

97. Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nature Reviews Cancer* **10**, 457–469 (2010).

98. Sanchez, R. & Zhou, M.-M. The PHD finger: a versatile epigenome reader. *Trends in Biochemical Sciences* **36**, 364–372 (2011).

99. Zeng, L. *et al.* Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**, 258–62 (2010).

100. Ramón-Maiques, S. *et al.* The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18993–8 (2007).

101. Wang, G. G. *et al.* Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* **459**, 847–51 (2009).

102. Wysocka, J., Reilly, P. T. & Herr, W. Loss of HCF-1–Chromatin Association Precedes Temperature-Induced Growth Arrest of tsBN67 Cells. *Molecular and Cellular Biology* **21**, 3820–3829 (2001).

103. Xie, L. *et al.* KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. *the The European Molecular Biology Organization Journal* **30**, 1473–1484 (2011).

104. Liefke, R. *et al.* Histone demethylase KDM5A is an integral part of the core Notch-RBP-J repressor complex. *Genes & Development* **24**, 590–601 (2010).

105. Barrett, A. *et al.* Breast cancer associated transcriptional repressor PLU-1/JARID1B interacts directly with histone deacetylases. *International journal of cancer. Journal international du cancer* **121**, 265–75 (2007).

106. Ernst, P., Mabon, M., Davidson, A. J., Zon, L. I. & Korsmeyer, S. J. An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Current Biology* **14**, 2063–2069 (2004).

107. Kuraoka, I. *et al.* Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *The Journal of Biological Chemistry* **283**, 940–950 (2008).

108. Zhou, W., Chen, H. & Zhang, L. The PcG protein hPc2 interacts with the N-terminus of histone demethylase JARID1B and acts as a transcriptional co-repressor. *BMB reports* **42**, 154–159 (2009).

109. Steinman, R. A. Cell cycle regulators and hematopoiesis. *Oncogene* **21**, 3403–3413 (2002).

110. Smith, E., Lin, C. & Shilatifard, A. The super elongation complex (SEC) and MLL in development and disease. *Genes & Development* **25**, 661–672 (2011).

111. Chang, M.-J. *et al.* Histone H3 lysine 79 methyltransferase Dot1 is required for immortalization by MLL oncogenes. *Cancer Research* **70**, 10234–10242 (2010).

112. Bernt, K. M. *et al.* MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).

113. Nguyen, A. T., Taranova, O., He, J. & Zhang, Y. DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis. *Blood* **117**, 6912–6922 (2011).

114. Jo, S. Y., Granowicz, E. M., Maillard, I., Thomas, D. & Hess, J. L. Requirement for Dot1l in murine postnatal hematopoiesis and leukemogenesis by MLL translocation. *Blood* **117**, 4759–4768 (2011).

115. Andersen, M. K., Christiansen, D. H. & Pedersen-Bjergaard, J. Amplification or duplication of chromosome band 21q22 with multiple copies of the AML1 gene and mutation of the TP53 gene in therapy-related MDS and AML. *Leukemia official journal of the Leukemia Society of America Leukemia Research Fund UK* **19**, 197–200 (2005).

116. Avet-Loiseau, H. *et al.* Amplification of the 11q23 region in acute myeloid leukemia. *Genes chromosomes cancer* **26**, 166–170 (1999).

117. Cuthbert, G. *et al.* MLL amplification in acute leukaemia: a United Kingdom Cancer Cytogenetics Group (UKCCG) study. *Leukemia official journal of the Leukemia Society of America Leukemia Research Fund UK* **14**, 1885–1891 (2000).

118. Wang JF, J, L., YF, X. & W, B. State of Chromosome 11q23 in T-ALL/LBL and their relation to prognosis. *Appl Immunohistochem Mol Morphol* **20**, 375–380 (2012).

VITA

The author, Gayathree Raman, was born to Ramasubramanian Raman and Bala Raman and was raised in Tirunelveli, India. She has a younger brother, Rajaram, who is a hydraulics engineer. She is married to Venkat Veeramani, Ph.D., and is blessed with an energetic two and half year old son, SreeJay.

Gayathree completed her high school education at Sri Jayendra Saraswathi Golden Jubilee School, Tirunelveli. She received her Bachelor's degree in Microbiology from Manonmaniam Sundaranar University, India (2000) and her Master degree in the field of Medical Laboratory Technology from Birla Institute of Technology & Science (BITS), India (2003). Her thesis was on identifying the genetic and epigenetic mechanisms that lead to retinoblastoma. She was awarded several meritorious awards including Gold Medal (2000) for securing first rank in her University, Best Performance in Clinical Genetics (2003) and Best Outgoing Student (2003) in M.S. Medical Laboratory Technology from BITS, India.

After graduating from BITS, Gayathree joined the Medical Research Foundation (India) as a Research Fellow and worked on single nucleotide polymorphisms in candidate genes and their association with diabetic retinopathy, molecular genetics of glaucoma and congenital cataract.

Gayathree joined the Molecular Biology Program of Loyola University Chicago in August 2006. She is currently completing her doctoral studies under the guidance of Dr. Manuel O. Diaz. Her research explores the biological function of MLL PHD3 binding to H3K4Me3 and Cyp33.

Gayathree has presented many posters, papers and published articles. She won an award for her poster at the Gordon Conference for Cancer Genetics and Epigenetics in 2011. She was provided an opportunity to write a proposal and mentor an undergraduate student, through the Research Mentoring Program (RMP) funded by Loyola University Chicago. She has also served as research mentor for high school, undergraduate and medical students. After completing her Ph.D., she intends to pursue research and teaching in the field of epigenetics.